

UNIVERSIDAD AUTÓNOMA DE MADRID  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE BIOLOGÍA MOLECULAR  
CENTRO DE BIOLOGÍA MOLECULAR “SEVERO OCHOA”

**MECANISMOS IMPLICADOS EN LA GENERACIÓN  
DE ISOFORMAS DE LA PROTEÍNA 4.1R**

Memoria presentada por  
**Eva Pilar Lospitao Ruiz**  
Para optar al grado de Doctor en Ciencias Biológicas

Directora de tesis:  
**Dra. Isabel Correas Hornero**  
Departamento de Biología Molecular  
Universidad Autónoma de Madrid  
Centro de Biología Molecular  
“Severo Ochoa” CSIC/UAM

**Madrid, 2007**

Esta tesis ha sido desarrollada en el centro de Biología Molecular “Severo Ochoa” bajo la dirección de la Dra. Isabel Correas Hornero. Para ello se contó con una beca de formación de personal investigador del Ministerio de Educación y Ciencia.



## AGRADECIMIENTOS.

Parecía que nunca iba a llegar este momento, pero como a la tercera va la vencida, al final llegó. Todo lo que he escrito aquí es fruto del esfuerzo de muchas personas que me apoyaron científica y personalmente.

A nivel científico tengo que agradecer a Isabel el haberme ofrecido la oportunidad de trabajar en su laboratorio y apoyarme durante estos cuatro años. Por supuesto no ha sido el único apoyo que he recibido por parte de la gente del CBM. En primer lugar la gente de mi laboratorio, el C011 de biológicas, mis chicas. Carmen, eres la persona más genial que me he encontrado en ciencia, ha sido un placer poder trabajar contigo y compartir tantos momentos, no sabes lo que te echo de menos. Me enseñaste que se puede ser mucho más, sin dejar de ser un buen científico. Altea, que difícil agradecerte tanto. Te lo he dicho muchas veces y es cierto: deberías ser coautora de esta tesis, pero eso no es nada en comparación con todo el apoyo personal y la amistad que siempre me has brindado. Muchísimas gracias por estar siempre que te necesité. Cristina, muchas gracias por tu ayuda, por todas esas risas y por animarme en muchos momentos. La otra chica del C011, la peque, Anita, quiero agradecerte todo lo que me has apoyado estos meses, tus ánimos, tus consejos y ese arrojo que le pones a todo que anima en cada momento a seguir trabajando con una sonrisa. A pesar de que ya no estén y de que a algunos no los he conocido, es necesario que agradezca el trabajo previo de las personas que estuvieron en el C011, entre ellos Carlos, Guillermo, Maria-José, Alicia, Eli, etc. Sin sus aportaciones mi trabajo nunca habría sido posible.

Quiero agradecer a todos los miembros del C210 su apoyo. Aunque pertenezcáis a otro laboratorio siempre me considerasteis una más, no sabéis cuanto me ayudó esto en momentos de "soledad". A Miguel, que siempre tuvo un hueco para escuchar mi trabajo y sugerirme millones de experimentos. A Juanfra, científicamente quiero agradecerte todo el apoyo que me has prestado en cada paso de esta tesis. En lo personal me faltaría espacio, pero muchas gracias por tener siempre un hueco para escuchar mis problemas, por tu sensibilidad y por la confianza que depositaste en mí. Eres un encanto. Ali (linfocita), ups, qué difícil. Creo que tengo que darte las gracias por ser como eres, tan vital. Nadie puede derrochar tanta energía como tú y además contagiarla. Muchas gracias por tu amistad Ali y por supuesto por toda la ayuda con los linfocitos. En el CBM se te echa mucho de menos. Olga, de cuántos apurillos me sacaste. Muchísimas gracias por tu ayuda con los *Northern*, por tus consejos. Laura, muchas gracias por tus ánimos en los momentos difíciles. A los otros chico del C-210 Ricardo, Alberto y más recientemente Fernando; muchísimas gracias por prestarme de todo, por escucharme y por animarme siempre. No puedo olvidarme de la gente que ahora mismo no está en el C-210 pero que me han ayudado en todo momento. A Noe por esa ironía combinada con un humor inmejorable en todo momento que siempre me animaron a ser un poco más optimista. Sergio más conocido como Machoman (no seré yo la que rompa la tradición de llamarte así), es envidiable la capacidad de trabajo, la serenidad que siempre demuestras, ojala la vida te depara lo mejor porque lo mereces sin duda. Maika, la persona más increíble que nunca he conocido. Es admirable tu sensibilidad y esa forma que tienes de ver las cosas como nadie más puede. Muchas gracias por todos tus consejos Maika, ojala vuelvas muy pronto al C210 porque se te echa de menos por aquí.

Quiero agradecer también a otras personas del CBM todo su apoyo. Muy especialmente toda la ayuda que me ha prestado Iván Ventoso. Gracias por tener siempre tiempo para escuchar mi trabajo, aconsejarme, darme millones de ideas geniales, y prestarme material. Gracias también a Virginia por su ayuda con el citómetro y a toda la gente del laboratorio C104 siempre dispuestos a ayudarme con la citometría de la que saben tanto. Quiero agradecer también su ayuda al servicio de microscopía del CBM, a Carlos, M<sup>a</sup> Ángeles, Teresa y Verónica; a la gente del servicio de informática, especialmente a Diego Díaz; a las chicas de seguridad biológica y a África que siempre tuvo un huequito para charlar y meterse conmigo. Gracias por hacerme reír siempre.

Existieron otros dos laboratorios que han participado en mi formación a pesar de que mi estancia en ellos fue previa al desarrollo de esta tesis. No puedo olvidarme de toda la gente que me enseñó en ellos ni de las personas que se convirtieron en algo más que compañeros de trabajo. Del C101 no puedo dejar de agradecer a los que yo llamo mis padres científicos todo su apoyo. Gracias Felipe, por todo. Por enseñarme tanto de “bichitos raros” y sobre todo por la confianza que siempre has depositado en mí, fue un placer trabajar contigo, y mucho más seguir viéndote de vez en cuando. Y cómo no a mi madre científica, Ángeles. Tengo que darte las gracias por esa forma de ser que tienes, por tu sinceridad apabullante, por tu amistad y apoyo, y sobre todo porque nunca me divertí tanto trabajando con nadie. De los comienzos tengo que dar también las gracias a Javi Chichón, compañero de fatigas, A Javi II, A Eva “la Mancheguita”, a Nuria I y II, a Irma, a Moustá, a Jose Luis, a Emiliano, a las Raqueles, muchas gracias a todos por ese magnífico comienzo en ciencia.

En cuanto a mi segundo laboratorio, en el Hospital Carlos III, no puedo más que agradecer a todos, científicos y médicos todo lo que hicieron por mí. En especial tengo que agradecer a África su apoyo incondicional en cada momento, su enorme capacidad de trabajo, y por encima de todo eso el haberse convertido en mi segunda mamá científica, por todos sus consejos y amistad. Fue un enorme placer tenerte de jefa. Del hospital tengo que agradecer enormemente el trato recibido por Vicente Soriano, así como por el resto de componentes de su equipo, entre ellos, Sylvina. ¡Qué argentina, madre mía!; Vicky, ojala siguiéramos cerca. Evita, qué entusiasmo y alegría le pones a cada experimento, a Mariola, a Maite, a Angélica, a todos... Os he echado mucho de menos estos cuatro años.

Por supuesto tengo que dar las gracias a mis amigos, los de siempre... Sabéis lo que me ha costado esto, así que creo que solo puedo deciros que si no hubieseis estado cerca no habría podido. Gracias Laura. Procura no cambiar, sigue como a los 15 porque ya entonces eras estupenda. A Dani, muchas gracias por estar siempre, incluso cuando no me entiendes y a Jose porque, aunque no lo creas, me encantan tus consejos y sí, a veces llevas razón; no sabéis como os echo de menos en estos momentos. A Raquel por sus ánimos, su ironía, su sonrisa, eres genial. A Cristina, por todos esos cafés, por mantener nuestra amistad. A Sando, por su ternura, su saber estar, por ser tan buen amigo, muchas gracias. Cómo no, a las mellis. A Olga (la melli morena) ¡qué chica!. Muchas gracias por estar cerca de mí y de los míos, ayudándonos tanto siempre. A Susi (la melli rubia) gracias por apoyarme siempre Susi y por cuidarme tanto... ¡No os merezco !

A pesar de que ocupen la parte final de estos agradecimientos mi familia nunca se ha quedado atrás en lo que a apoyo se refiere. Quiero dar las gracias a mis tíos y primos que siempre han estado cerca apoyándome y preguntándome por “mis bichitos” como dicen ellos. A mi Prima Laura por su ayuda con el inglés y sus ánimos. A Carlitos, el único niño de la familia por hacerme reír tantas veces. A mi tía por todo su cariño y por esos desayunos. Y, por supuesto, a mi tío “El de los poderes” que siempre está dispuesto a ayudar. También quiero dar las gracias a alguien que a pesar de haberse ido hace mucho sigue conmigo, mi abuelo, estoy segura que le habría gustado leer que él también contribuyó a que su *Música* acabara esta tesis. A los contrarios, Nines, La Roter, De la Moto, Aitor, Ivanna, a la yaya... muchísimas gracias por el apoyo y los ánimos que me habéis prestado durante estos meses.

Ahora viene lo difícil porque tengo que dar las gracias a cuatro personas importantísimas en mi vida, mi pilar, que me lo han dado todo y siempre. Este trabajo nunca se habría llevado a cabo sin su apoyo, así que no bastará con lo que aquí pueda escribir para agradecerles todo lo que han hecho por mí. Papa, gracias por enseñarme que no merece la pena perder tiempo en lamentaciones, gracias por inculcarme ese arrojo que te caracteriza, esa fuerza y todo ese cariño. Mama, necesitaría tantas páginas... no caben las palabras para agradecer toda tu paciencia conmigo, los ánimos, cariño y todos esos buenos consejos. Sin vosotros dos este trabajo nunca habría sido posible, espero que os guste leerlo. Sois los mejores padres del mundo. Lulu, muchas gracias por aconsejarme siempre, por haberme apoyado tanto a pesar de que a veces es complicado entender todo lo que rodea a la ciencia. Gracias por todo tu cariño, por tu carácter, por tus juegos, por todos esos momentos. Bárbara... Dicen que lo bueno se hace esperar, seguro

que por eso eres la pequeña. Quiero darte las gracias por todo lo que me has ayudado y enseñado, por tus ánimos, especialmente este mes encerrada en casa, por tu calma y esa serenidad, admirable en todo momento. Te quiero *xota*.

Ya sólo me queda agradecerle a una persona. Te dejé al final Pablo. No sé muy bien cómo agradecerte la paciencia que tienes conmigo, esa calma que me infundes. Gracias por aprender tanto de traducción como de construcción, por apoyarme siempre a pesar de que esto de la ciencia no sea fácil y te caiga muy lejos. Gracias por tu cariño. Sin ti esto habría sido un imposible.

A alguien que me llamaba *Música*, mi abuelo.

A mis padres y mis hermanas.

A Pablo.

# Índice



<b><u>SUMMARY</u></b>	<b>1</b>
<b><u>ABREVIATURAS</u></b>	<b>2</b>
<b><u>INTRODUCCIÓN</u></b>	<b>5</b>
<b>1. Mecanismos generadores de diversidad proteica a partir de un único gen.</b>	<b>6</b>
1.1. Mecanismos que actúan a nivel de la transcripción o maduración del ARN.	6
1.1.1. Promotores alternativos.	7
1.1.2. Sitios de poliadenilación.	7
1.1.3. Procesamiento alternativo del pre-ARN.	9
1.1.4. Edición del ARN.	11
1.2. Mecanismos que actúan al inicio de la traducción.	12
1.2.1. El modelo del rastreo del ribosoma ( <i>scanning</i> ).	12
1.2.2. <i>Leaky scanning</i> .	15
1.2.3. Reiniciación.	16
1.2.4. <i>Shunting</i> (Salto del ribosoma).	17
1.2.5. Sitios de entrada interna del ribosoma (IRES).	17
<b>2. Proteína 4.1R: un miembro de la superfamilia 4.1.</b>	<b>19</b>
2.1.1. La proteína 4.1R.	20
2.1.2. <i>EPB41</i> : el gen codificante para 4.1R.	21
2.1.3. Procesamiento alternativo del pre-ARNm de 4.1R.	21
2.1.4. Dominios estructurales de la proteína 4.1R	23
2.1.5. Funciones de la proteína 4.1R en el eritrocito.	25
2.1.6. Funciones de la proteína 4.1R en células nucleadas.	26
<b><u>OBJETIVOS</u></b>	<b>28</b>
<b><u>MATERIALES Y MÉTODOS</u></b>	<b>30</b>
<b>1. Cultivos celulares</b>	<b>31</b>
1.1. Líneas celulares de mamíferos.	31
1.2. Cepas bacterianas.	32
<b>2. Anticuerpos</b>	<b>32</b>
2.1. Anticuerpos primarios.	32
2.2. Anticuerpos secundarios	33
<b>3. Vectores de expresión</b>	<b>33</b>
<b>4. Oligonucleótidos</b>	<b>34</b>
<b>5. Enzimas</b>	<b>35</b>
5.1. Enzimas de restricción.	35
5.2. Enzimas de modificación del ADN, Polimerasas y Retrotranscriptasas.	35

<b>6. Generación de proteínas truncadas, mutadas y de fusión de 4.1R.</b>	<b>35</b>
6.1. Construcciones de ADNc parciales y quimeras de 4.1R	35
6.2. Mutagénesis del sitio de iniciación de la traducción en el exón 4 (ATG2)	36
6.3. Mutagénesis que introduce un codón de parada de la traducción en el exón 2.	36
6.4. Construcción bicistrónica DsRed-4.1-EGFP.	36
7. Secuenciación de los ADNc.	37
8. Transcripción y traducción <i>in vitro</i> .	37
9. Transfección transitoria de ADNc en células de mamífero.	37
10. Obtención de extractos proteicos.	38
11. Aislamiento de centrosomas de células MOLT-4.	38
12. Electroforesis de proteínas (SDS-PAGE) y electrotransferencia ( <i>Western Blot</i> ).	38
13. Microscopía.	39
14. Generación de anticuerpos policlonales.	39
15. Inmunoprecipitaciones.	39
16. Centrifugación por velocidad de sedimentación en gradientes de densidad en sacarosa.	40
17. Aislamiento de ARN total de células en cultivo.	40
18. Transferencia del ARN a membrana e hibridación ( <i>Northern Blot</i> ).	40
19. Aislamiento y clonación de los ADNc para isoformas de 4.1R sin dominio CTD.	41
19.1. Retrotranscripción.	41
19.2. Reacción en cadena de la polimerasa.	41
19.3. Clonación de los ADNc amplificados.	41
19.4. Subclonación de los ADNc.	41
20. Experimentos de citometría de flujo.	42
<b><u>RESULTADOS</u></b>	<b>43</b>
<b>1. Generación de dos isoformas de 4.1R desde los ADNc que contienen el exón 2'</b>	<b>44</b>
1.1. Algunos ADNc para 4.1R <sup>135</sup> originan dos proteínas <i>in vitro</i> .	44
1.2. Las proteínas de menor tamaño sintetizadas <i>in vitro</i> desde ADNc para 4.1R <sup>135</sup> presentan el mismo tamaño que las isoformas correspondientes de 4.1R <sup>80</sup> .	45
1.3. Las proteínas de menor tamaño obtenidas a partir de ADNc de 4.1R <sup>135</sup> son isoformas traducidas desde el ATG2.	46
1.4. Las isoformas de 4.1R <sup>80</sup> obtenidas a partir de ADNc para 4.1R <sup>135</sup> no son consecuencia de una proteólisis.	47
1.5. La región 5' anterior al ATG2 de 4.1R es suficiente para dirigir la síntesis de la proteína 4.1R <sup>80</sup> .	47
1.6. Expresión <i>in vivo</i> de ADNc para isoformas de 4.1R <sup>135</sup> .	49
1.7. Análisis de la expresión <i>in vivo</i> de diferentes isoformas de 4.1R <sup>135</sup> mediante inmunofluorescencia.	50
1.8. La isoforma de 4.1R <sup>80</sup> generada <i>in vivo</i> desde el ATG2 no es un producto de proteólisis.	51
1.9. La secuencia anterior al ATG2 del exón 4 genera la isoforma de peso molecular intermedio <i>in vivo</i> .	52
1.10. No existe un promotor críptico en ADNc para 4.1R <sup>135</sup> .	53
1.11. Generación de un vector bicistrónico.	54
1.12. Expresión del vector bicistrónico en células COS-7.	55
1.13. Relación EGFP/DsRed en diferentes líneas celulares.	56
1.14. La producción de EGFP no es consecuencia de un procesamiento del ARNm.	57
1.15. La secuencia de 4.1R clonada en el vector bicistrónico determina la eficiencia de producción del segundo cistón.	58

1.16. Los ADNc para 4.1R <sup>135</sup> con exón 4 contienen una secuencia de entrada directa para el ribosoma (IRES).	59
1.17. Comparación de la secuencia comprendida entre el ATG1 y ATG2 del gen que codifica para 4.1R en diversas especies.	60
1.18. Análisis de la traducción del segundo cistron en células en cultivo sometidas a distintos tratamientos.	63
1.19. Las dos isoformas de 4.1R generadas desde un único ADNc no interaccionan.	63
<b>2. Aislamiento de nuevas isoformas de 4.1R que carecen del dominio carboxilo terminal (CTD).</b>	<b>64</b>
2.1. Características del clon BC039079.	65
2.2. Expresión del clon BC039079.	65
2.3. Nuevas isoformas sin dominio CTD en cerebro humano.	67
2.4. Composición exónica de los ADNc aislados.	68
2.5. Expresión de los ADNc para isoformas de 4.1R sin el dominio CTD	70
2.6. Localización intracelular de las isoformas de 4.1R sin dominio CTD.	72
2.7. Generación de un anticuerpo específico para las isoformas de 4.1R sin dominio CTD.	72
2.8. Análisis de la expresión endógena de las isoformas 4.1R sin CTD en distintos tipos celulares.	74
2.9. Localización de las isoformas de 4.1R sin CTD en diferentes líneas celulares.	75
2.9.1. Localización en células interfásicas.	75
2.9.2. Localización en células en división.	78
2.10. Las isoformas de 4.1R sin CTD colocalizan con aurora B y con anexina II en células en citocinesis.	80
2.11. Las isoformas de 4.1R sin dominio CTD se localizan en el centrosoma de células COS-7 y MOLT-4.	80

## **DISCUSIÓN** **85**

<b>1. Un mecanismo de traducción interna permite la generación de isoformas de peso molecular intermedio de 4.1R.</b>	<b>86</b>
1.1. Dos isoformas de 4.1R son expresadas desde un único ADNc.	86
1.2. Una traducción interna está implicada en la generación de la isoforma de 4.1R de peso molecular intermedio.	87
1.3. Traducción interna. <i>Leaky scanning</i> , <i>Shunting</i> o IRES.	87
1.4. La traducción interna del ADNc para 4.1R es dependiente de la línea celular y específica de la secuencia del ADNc.	88
1.5. Posibles implicaciones biológicas de la existencia de un IRES en la secuencia del ADNc para 4.1R.	90
1.6. El IRES de 4.1R se encuentra en la región codificante.	91
<b>2. Isoformas de 4.1R sin dominio CTD.</b>	<b>92</b>
2.1. Nuevo patrón de procesamiento en el pre-ARNm de 4.1R.	92
2.2. Composición exónica de los ADNc para isoformas de 4.1R sin CTD.	93
2.3. Localización de las isoformas de 4.1R sin dominio CTD.	93
<b>3. Los mecanismos que generan isoformas de 4.1R actúan a distintos niveles.</b>	<b>95</b>

## **CONCLUSIONES** **96**

## **BIBLIOGRAFÍA** **98**

## **ANEXO** **107**

## Summary

## SUMMARY

Red blood cell protein 4.1, 4.1R, is an extreme variation on the theme of isoform multiplicity and involvement of alternative pre-mRNA splicing events for the generation of 4.1R protein diversity has been widely explored. Two types of 4.1R isoforms varying in N-terminal extensions are originated by alternative splicing events involving exon 2'. This exon encompasses translation initiation site AUG1 and hence 4.1R mRNAs containing exon 2' regulate expression of longer (135 kDa) isoforms of 4.1R protein. By contrast, 4.1R mRNAs excluding exon 2' allow expression of shorter (80 kDa) isoforms, the synthesis of which is initiated at the AUG2 translation initiation site comprised in exon 4. The current study reports that 4.1R mRNAs containing exons 2' and 4 regulate expression of both longer and shorter 4.1R isoforms. Analysis of *in vitro* expression of a set of 4.1R cDNAs containing exon 2' showed an unexpected result: that two proteins 4.1R, longer (~135-kDa) and shorter (~80-kDa), were synthesized. Mutational studies indicated that the shorter protein 4.1R was not a proteolytic product of the longer one but a product synthesized from the downstream AUG2 site. Results of further experiments showed that the sequence 5'-upstream of exon 4 was essential for the use of the AUG2 as internal initiation site of translation. When this sequence was introduced in a bicistronic vector, it directed the synthesis of the second cistron even when a hairpin structure was added 5'-upstream of the first cistron and only one mRNA species was detected by Northern blot analysis. *In vivo* expression of this set of 4.1R cDNAs confirmed that shorter and longer 4.1R isoforms were generated but not when the 4.1R sequence was introduced in a promoterless vector. These results indicate that the 5' region upstream of exon 4 contains an internal ribosome entry site (IRES) element which directs the synthesis of 80-kDa 4.1R isoforms from mRNAs containing exon 2'. These data show that generation of 4.1R isoforms is also regulated at the translational level.

The current study identified a set of 4.1R cDNAs coding a new group of isoforms lacking the carboxy-terminal domain (CTD) characteristic of all 4.1R proteins identified so far. An alternative poly-A signal present in the sequence between exons 17 and 18 of the *EPB41* gene and the inclusion of twenty seven nucleotides comprising a stop codon regulated the production of this set of 4.1R isoforms lacking the CTD. Generation of an antiserum that specifically recognized this set of isoforms as well as transfection experiments using the isolated 4.1R cDNAs have allowed to determine the localization of these isoforms within the cell.

Together, our data implicate two mechanisms not previously described as involved in the regulation of 4.1R expression: internal translation and alternative polyadenylation with the inclusion of a premature stop codon.

## Abreviaturas

**ABREVIATURAS Y TÉRMINOS EN INGLÉS**

<b>ADN</b>	Ácido desoxirribonucleico.
<b>4.1R<sup>135</sup></b>	Grupo de isoformas de 4.1R de peso molecular alto.
<b>4.1R<sup>80</sup></b>	Grupo de isoformas de 4.1R de peso molecular intermedio.
<b>A</b>	Adenina.
<b>Aa</b>	Aminoácido.
<b>ADAR</b>	<i>Adenosine Deaminase Active on RNA</i> , desaminasa del ARN.
<b>ADNc</b>	ADN complementario, obtenido por retrotranscripción de ARNm.
<b>ARN</b>	Ácido Ribonucleico.
<b>ARNm</b>	ARN mensajero.
<b>ATP</b>	Adenosín trifosfato.
<b>BBP</b>	<i>Branch point Binding Protein</i> . Proteína de unión al punto de bifurcación.
<b>C</b>	Citosina.
<b>CAP</b>	Caperuza constituida por una guanosina metilada que se añade al extremo 5' del ARNm maduro.
<b>CMV</b>	Citomegalovirus.
<b>CN</b>	Control negativo.
<b>CTD</b>	Dominio carboxilo terminal de las isoformas 4.1R.
<b>DAPI</b>	4', 6'-diamino-2-fenilindol.
<b>DMSO</b>	Dimetil sulfóxido.
<b>DsRed</b>	<i>Discosoma Red fluorescence Protein</i> , Proteína roja fluorescente de <i>Discosoma</i> .
<b>EDTA</b>	Ácido etilen diamino tetraacético.
<b>EGFP</b>	<i>Enhanced Green Fluorescence Protein</i> . Proteína verde fluorescente.
<b>EPB41</b>	Gen que codifica para la proteína 4.1R.
<b>ERM</b>	Ezrina, Radixina y Moesina; familia de proteínas pertenecientes a la superfamilia de 4.1.
<b>FACS</b>	<i>Fluorescence Activated Cell Sorting</i> , separación de células fluorescentes mediante citometría de flujo.
<b>FERM</b>	<i>Four point one, Ezrin, Radixin and Moesin</i> , dominio proteico que caracteriza a la superfamilia de proteínas de la que forman parte las ERM y la proteína 4.1.
<b>Fig</b>	Figura.
<b>G</b>	Guanina.
<b>g</b>	Gramo.
<b>G1</b>	Fase del ciclo celular que transcurre entre el fin de una mitosis y el comienzo de la síntesis de ADN.
<b>G2</b>	Fase del ciclo celular que transcurre entre el fin de la síntesis de ADN y el comienzo de una mitosis.
<b>GPC</b>	Glicoforina C.
<b>HP</b>	Dominio <i>Head Piece</i> . Dominio de 209 aminoácidos presentes en las isoformas de 4.1R de peso molecular alto.
<b>IRES</b>	<i>Internal ribosome entry site</i> . Secuencia que dirige la entrada directa del ribosoma.
<b>ITAFs</b>	<i>Internal trans-acting factors</i> . Factores transactivadores internos.
<b>KCal</b>	Kilocaloría.

<b>kDa</b>	Kilodalton.
<b>M</b>	Molar.
<b>m7Gppp</b>	7-metilguanosina trifosfato.
<b>MDCK</b>	<i>Madin-Darby Canine Kidney</i> , línea celular establecida.
<b>Met-tRNA</b>	Metionil ARN de transferencia.
<b>MGC</b>	<i>Mammalian Gene Collection</i> , Colección de Genes de Mamíferos.
<b>ml</b>	Mililitro.
<b>mM</b>	Milimolar.
<b>PADP</b>	Proteína de unión a poliadenilato.
<b>pb</b>	Pares de bases.
<b>PBS</b>	Solución salina tamponada con fosfatos.
<b>PCR</b>	Reacción en cadena de la polimerasa.
<b>PI</b>	Punto isoeléctrico.
<b>Pm</b>	Peso molecular.
<b>Pre-ARN</b>	ARN mensajero no procesado.
<b>R</b>	Adenina o guanina.
<b>RT-PCR</b>	Retrotranscripción y reacción en cadena de la polimerasa.
<b>S</b>	Sveldbergs.
<b>SABD</b>	<i>Spectrin Actin Binding Domain</i> (Dominio de unión a actina y espectrina de la proteína 4.1).
<b>SDS-PAGE</b>	Separación de proteínas por electroforesis en geles de poliacrilamida y en presencia de SDS.
<b>snRNAs</b>	<i>Small nuclear RNAs</i> , ARN nucleares pequeños.
<b>snRNPs</b>	Small nuclear ribonucleoproteins, ribonucleoproteínas nucleares pequeñas.
<b>T</b>	Timina.
<b>ARNt</b>	ARN de transferencia.
<b>U</b>	Uracilo.
<b>UTR</b>	<i>Untranslated Region</i> . Región no traducida del RNAm.
<b>Y</b>	Citosina o Uracilo.
<b>µg</b>	Microgramo.

Términos en inglés:

<b>Knockout</b>	Inactivación total o parcial de un gen llevada a cabo mediante técnicas de ADN recombinante.
<b>Leaky scanning</b>	Rastreo del ribosoma sobre el ARNm que permite el reconocimiento de un triplete iniciador posterior a otro que también se utiliza como codón iniciador.
<b>Scanning model</b>	Modelo de inicio de la traducción mediante el rastreo del ribosoma.
<b>Shunting</b>	Salto del ribosoma. Cuando el ribosoma está rastreando el ARNm en busca del triplete iniciador de la traducción puede no rastrear una zona de éste, la cual "salta".
<b>Western Blot</b>	Transferencia de proteínas a membrana y detección mediante anticuerpos.
<b>Northern Blot</b>	Transferencia de RNA a membrana e hibridación con sonda marcada.



# Introducción

## INTRODUCCIÓN

El análisis del genoma humano ha permitido establecer que el número de genes que codifica para las proteínas identificadas es mucho menor de lo inicialmente estimado (Southan, 2004). Así, a pesar de que se estimó la existencia de 150.000 genes, se sabe en la actualidad que este número está comprendido entre 20.000 y 25.000 (Hattori, 2005). Comparando éste con el número de genes de organismos como *Drosophila melanogaster* (13.601), *Caenorhabditis elegans* (18.424) o *Arabidopsis thaliana* (27.000) se puede concluir que no existe un aumento en el número de genes de especies más complejas que justifiquen dicha complejidad (Touriol et al., 2003). Los 25.000 genes que componen el genoma humano son capaces de codificar unas 100.000 proteínas, lo que necesariamente implica que diferentes proteínas son codificadas por un único gen. Se conocen muchos de los mecanismos que permiten la producción de diferentes proteínas a partir de un único gen. Sin embargo, existe menos información de cómo se regulan muchos de estos mecanismos y la relación que mantienen entre sí. En muchos casos se desconoce todavía cuántas proteínas genera un gen concreto, ya que los diferentes tipos de expresión de un gen dependen del organismo, el tejido e incluso la célula y de su entorno. Esto hace que el estudio de la generación de diversidad proteica sea mucho más complejo que el estudio del genoma en sí mismo. En este trabajo abordaremos el estudio de dos estrategias que permiten la generación de diferentes isoformas de la proteína 4.1R. Se describen a continuación algunos mecanismos mediante los cuales se pueden generar dos proteínas a partir de un único gen y posteriormente se exponen los datos sobre la proteína 4.1R y el gen que codifica para ella.

### **1. Mecanismos generadores de diversidad proteica a partir de un único gen.**

Los diferentes mecanismos a través de los cuales se pueden generar diversas proteínas a partir de la información codificada en un gen actúan a distintos niveles. El primer nivel lo constituyen los mecanismos de recombinación génica como la meiosis y los procesos de recombinación génica que tienen lugar en los linfocitos. Existen mecanismos que actúan posteriormente, una vez que el ADN genómico se transcribe generando moléculas de ARN. Son estos últimos mecanismos en los que nos centraremos en la introducción de este trabajo.

#### **1.1. Mecanismos que actúan a nivel de la transcripción o maduración del ARN.**

La secuencia de ADN que contiene un gen se transcribe a ARN, éste es procesado y transportado al citoplasma celular para su posterior traducción a proteína. Una unidad simple de traducción en eucariotas (cistrón), se define como el molde de la polimerasa del ARN que contiene la información para un único producto proteico. Sin embargo, se sabe que en la mayoría de los casos un gen puede dar lugar a múltiples ARNm maduros mediante cuatro mecanismos diferentes que describiremos a continuación.

### 1.1.1. Promotores alternativos.

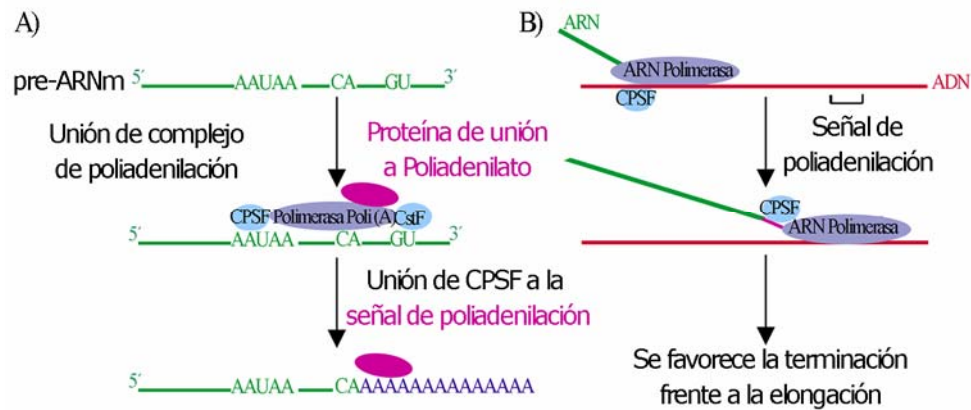
La transcripción es el proceso mediante el cuál la enzima ARN polimerasa copia la secuencia de un ADN a ARN. El inicio de la transcripción es un proceso que implica la formación del complejo de preiniciación en el promotor correcto y el momento adecuado, requisitos indispensables para la correcta síntesis de un pre-ARNm concreto. En este proceso intervienen diferentes factores de transcripción, de remodelación de la cromatina y elementos de unión al promotor. Los organismos pluricelulares han desarrollado una serie de componentes de la maquinaria de transcripción que permiten el desarrollo de programas de expresión génica complejos. Muchos de ellos intervienen en el reconocimiento y unión de elementos al núcleo del promotor (como la caja TATA). (Hochheimer and Tjian, 2003).

Existen muchos genes que contienen más de un promotor y que, por lo tanto, pueden generar diferentes proteínas en función del promotor utilizado para la transcripción. El uso de promotores alternativos da lugar al menos a dos ARNm con una región 5' diferente. Si en esta región 5' diferente existen codones para la iniciación de la traducción, se pueden codificar proteínas diferentes (Ayoubi and Van De Ven, 1996). Un ejemplo de gen con diferentes promotores es el gen que codifica para la distrofina. Este gen contiene 8 promotores alternativos, generando 8 proteínas diferentes en función del tejido en el que éstas se expresen (Strachan, 2003).

Se sabe que existe una estrecha relación entre la transcripción y el procesamiento alternativo del ARNm (Maniatis and Reed, 2002), así en muchos casos el uso de promotores alternativos está asociado a un patrón de procesamiento alternativo diferencial (Maniatis and Reed, 2002).

### 1.1.2. Sitios de poliadenilación.

Todos los ARNm eucariotas poseen en su extremo 3' unos 250 residuos de adenina, denominados cola de poli-A. Ésta interviene en fomentar la estabilidad del ARN y estimular la traducción (Mangus et al., 2003). Estos residuos son añadidos por la polimerasa poli(A). Una señal de 20-30 nucleótidos presentes en la zona 5' anterior al lugar de poliadenilación dirige este proceso. Esta señal consiste en la secuencia conservada 5'-AAUAAA-3'. Además, en el sitio en el que se añadirá la cola de poli-A se encuentran los nucleótidos 5'-CA-3' y 10-20 nucleótidos después la secuencia 5'-GU-3' (fig. 1A).



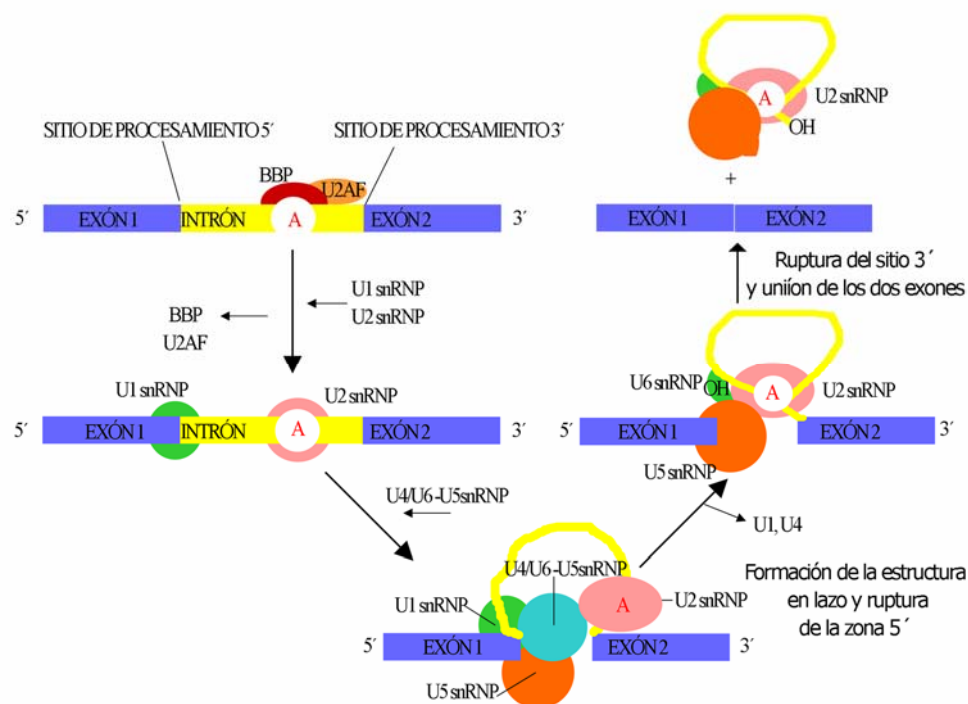
**Figura 1. Poliadenilación.** A) Esquema de las secuencias que intervienen en la unión del complejo de poliadenilación. Las proteínas CPSF y CstF se unen a la secuencia 5'-AAUAAA-3' y a secuencias ricas en GU, respectivamente. La secuencia rica en GU ha sido representada con un único GU. B) Esquema de la relación existente entre la poliadenilación y la terminación de la transcripción por la ARN polimerasa II. La proteína CPSF se une al complejo de transcripción hasta que se transcribe la secuencia AAUAAA, momento en el que se produce un cambio conformacional que libera a CPSF de la ARN polimerasa II y se favorece la terminación frente a la elongación.

En este proceso intervienen diferentes complejos proteicos, el factor de corte y poliadenilación (CPSF), el factor estimulante de la poliadenilación (CstF), la polimerasa poli(A) y la proteína de unión a poliadenilato (PADP) (Wahle and Rueggsegger, 1999). Esta última ayuda a la polimerasa a añadir adenosinas y parece influir en el tamaño de la cola de poli-A que se añade, así como en el mantenimiento de ésta una vez finalizada su síntesis. CPSF es reclutado al complejo de la polimerasa al inicio de la transcripción. Se mantiene unido mientras la polimerasa transcribe y se une a la señal de poli-A en cuanto ésta ha sido transcrita, momento en el que comienza la reacción de poliadenilación. Tanto CPSF como CstF parecen cambiar el contacto que mantenían con la polimerasa cuando ésta transcribe la señal de poli-A. Este cambio altera las propiedades del complejo de elongación, de forma que se favorece la terminación (fig. 1B). Al menos un 50% de los genes codificantes que posee el genoma de un mamífero presentan distintos sitios de poliadenilación, lo que puede generar la existencia de ARNm con extremos 3' diferentes (Ara et al., 2006). Los sitios de poliadenilación alternativos se suelen clasificar en función de que se encuentren en tándem en el mismo exón 3' o bien se encuentren en distintos exones 3' (Edwards-Gilbert et al., 1997). En este último caso, el procesamiento alternativo específico afecta a las señales de adenilación que se utilizarán. Esto último ocurre en el gen que codifica para la ATPasa transportadora de  $\text{Ca}^{2+}$  del retículo endoplásmico (SERCAs). Uno de sus pre-ARN mensajeros da lugar a dos proteínas que difieren en su dominio carboxilo debido a este mecanismo (Strehler and Zacharias, 2001).

Otro ejemplo lo constituye el gen de la calcitonina que da lugar a dos isoformas proteicas diferentes según el tejido en el que se expresen, gracias a una combinación de procesamiento y poliadenilación alternativos (Lou et al., 1995).

### 1.1.3. Procesamiento alternativo del pre-ARN.

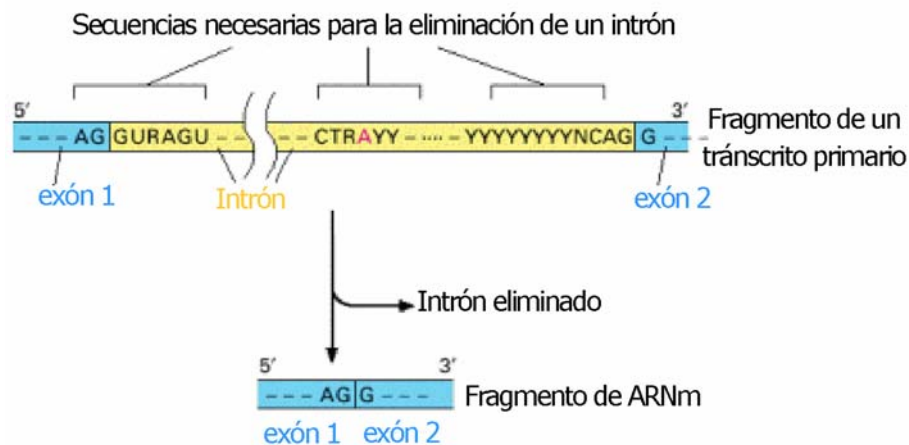
La secuencia codificante que contiene un gen eucariota está interrumpida por secuencias no codificantes denominadas intrones (Gilbert, 1978). Los fragmentos de secuencia codificantes se denominan exones. Tanto los exones como los intrones son transcritos por la polimerasa del ARN.



**Figura 2. Procesamiento alternativo del pre-ARN.** El procesamiento alternativo del ARN es catalizado por la unión de snRNPs (círculos coloreados) a diferentes proteínas (no mostradas) que conjuntamente forman el espliceosoma. Este reconoce señales de procesamiento en el pre-ARNm y une los dos extremos del intrón, permitiendo su escisión en dos pasos. El sitio de bifurcación es reconocido por la proteína de unión al punto de bifurcación (BBP) y por U2AF. En los siguientes pasos el snRNP U2 desplaza a BBP y U2AF y se une a la secuencia del punto de bifurcación. El snRNP U1 se une al extremo 5' de procesamiento. En ese momento el triple snRNP U4/U6-U5 se une al espliceosoma. Ocurren reorganizaciones en el ARN que crean un sitio activo en el espliceosoma y le posicionan en el lugar en el que ocurrirá el procesamiento del intrón.

El proceso mediante el cual se elimina los intrones del ARN recién transcrito o pre-ARN se conoce con el nombre de procesamiento alternativo (fig. 2) (Norton, 1994). La eliminación de cada intrón conlleva dos reacciones de transesterificación que permiten la unión

de dos exones, liberando el intrón. La maquinaria que cataliza este proceso es muy compleja, y está constituida por cinco ARN conocidos como *snRNAs* (*small nuclear RNAs*). Cada uno de ellos forma un complejo con, al menos, siete unidades proteicas formando las *snRNPs* (*small nuclear ribonucleoproteins*). Estos *snRNPs* forman el espliceosoma, encargado del procesamiento alternativo del pre-ARN. El espliceosoma es una maquinaria dinámica que reconoce el extremo 3' del exón anterior al intrón que será eliminado y la zona 5' del exón que le sigue. Existen tres secuencias que intervienen en la eliminación del intrón, el sitio 5' de ajuste, el sitio 3' y el punto de bifurcación (fig. 3) (Maniatis and Tasic, 2002). Este último se encuentra en la secuencia del intrón que formará la base de una estructura en lazo que permite al espliceosoma cortar el intrón. A pesar de que estos tres sitios tienen una secuencia consenso similar en todos los intrones, presentan una variabilidad alta lo que dificulta conocer los intrones dentro de un gen. Así, la elección de los sitios donde ocurrirá el procesamiento de un ARNm específico depende de varios factores, entre los que se encuentran las secuencias que acabamos de definir. Esto permite que una célula incluya o excluya determinadas secuencias de un pre-ARN dando lugar a diferentes ARN mensajeros maduros.



**Figura 3. Secuencias nucleotídicas consenso para la eliminación de intrones humanos.** Únicamente se requieren las tres zonas de nucleótidos señaladas con una llave. Se representa con A, G, U y C los nucleótidos estándar en el ARN; con una R se representan los nucleótidos A o G y con Y se representan los nucleótidos C o U. La A en rojo es la que se encuentra en el punto de bifurcación en el que se formará la estructura en lazo. Únicamente la secuencia GU, al comienzo del intrón, y la secuencia AG, al final de éste, están conservadas en las regiones consenso.

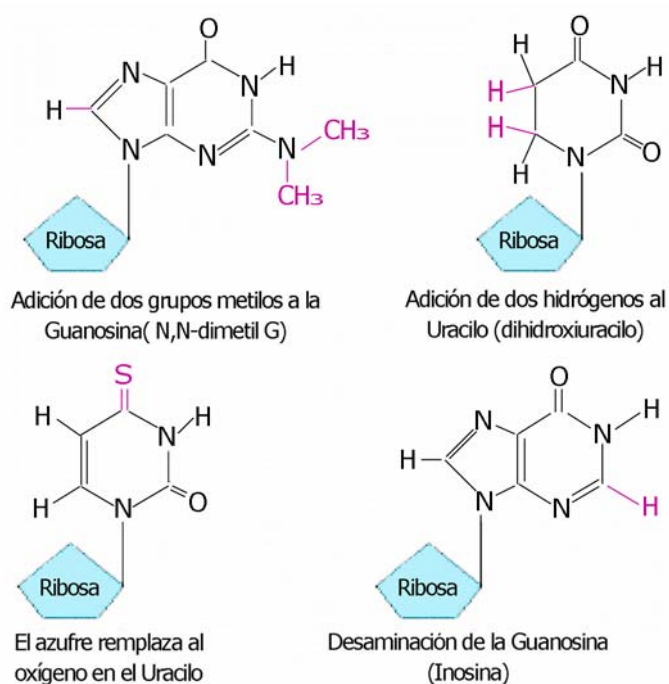
Las secuencias que se eliminan siempre se conocen como intrones mientras que secuencias que se eliminan en algunos tipos celulares o en algunas condiciones se denominan exones alternativos. Exones constitutivos son aquellos que no se eliminan. Existen factores activadores del procesamiento y factores inhibidores del procesamiento alternativo que contribuyen a la elección de los exones

alternativos que se incluirán en un ARNm determinado (Smith and Valcarcel, 2000). De esta forma, un gen concreto puede expresar diferentes combinaciones de exones en función del entorno en el que se exprese. Se estima que entre un 35 y un 60 % de los genes que dan lugar a varios transcritos lo hacen gracias a un procesamiento alternativo (Woodley and Valcarcel, 2002), y que un 40-60% de los genes del genoma humano sufren este tipo de procesamiento (Modrek and Lee, 2002).

Un ejemplo lo constituye el gen de la tropomiosina, proteína implicada en la contracción de las células musculares. Este gen puede generar diferentes ARNm que originan diferentes isoformas de esta proteína, algunas específicas de tejido (Gunning et al., 2005).

#### 1.1.4. Edición del ARN.

La edición del ARN consiste en la inserción o alteración de nucleótidos individuales de una molécula de ARN. En la figura 4 se muestran algunos de los cambios que pueden sufrir nucleótidos concretos.



**Figura 4. Edición del ARNm.** Se muestran algunos ejemplos de nucleótidos que sufren modificaciones covalentes dando lugar a nucleótidos inusuales en el ARNm.

El ejemplo más destacado de este proceso se ha observado en los transcritos que codifican para proteínas mitocondriales en tripanosomas y en plantas. Este proceso parece estar más limitado en mamíferos. El caso más representativo es la desaminación que sufren algunas adeninas convirtiéndose en inosina en algunos pre-ARNs. En algunos casos, esta modificación cambia el patrón de procesamiento alternativo o el significado de los tripletes. La inosina se empareja con



citosina, por lo que un cambio de adenina a inosina podría dar lugar a una proteína en la que ha cambiado un aminoácido. La enzima encargada de este proceso es la adenosina desaminasa (ADAR, *Adenosine Deaminase Active on RNA*)(Schaub and Keller, 2002). Se conocen también cambios de citosinas a uracilo en pre-ARN de mamíferos mediante este mecanismo (Wedekind et al., 2003). Un ejemplo lo presenta el ARNm para la apolipoproteína B de determinados tejidos, en el que una citosina es editada a uracilo generando un codón de parada de la traducción, que da lugar a una proteína truncada.

## 1.2. Mecanismos que actúan al inicio de la traducción.

La traducción es el proceso mediante el cual se genera un polipéptido a partir de un ARN mensajero maduro. Para que tenga lugar son necesarios el ARNm maduro, el ribosoma, diferentes ARN (ARNt) de transferencia y aminoácidos, así como un gran número de factores proteicos que conforman complejos que regulan y hacen posible este proceso. La traducción consta de tres etapas: el inicio, en el que el ribosoma ha de reconocer y anclarse al ARNm encontrando el triplete de iniciación; la elongación, durante la cual la secuencia de ARN será leída por los ribosomas, interpretando cada tres nucleótidos como un aminoácido concreto que será incorporado a la cadena polipeptídica; y la terminación, etapa en la que el ribosoma se libera del ARNm tras haber formado una proteína.

Cada secuencia de ARNm puede ser leída siguiendo tres posibles pautas de lectura en función del sitio en el que comience este proceso de descodificación. Por lo tanto, el lugar en el que se inicia la traducción determinará la proteína que va a codificar ese ARNm. Además, la etapa de la iniciación es la más regulada en el proceso de traducción. Se conocen distintos tipos de iniciación que se detallan a continuación.

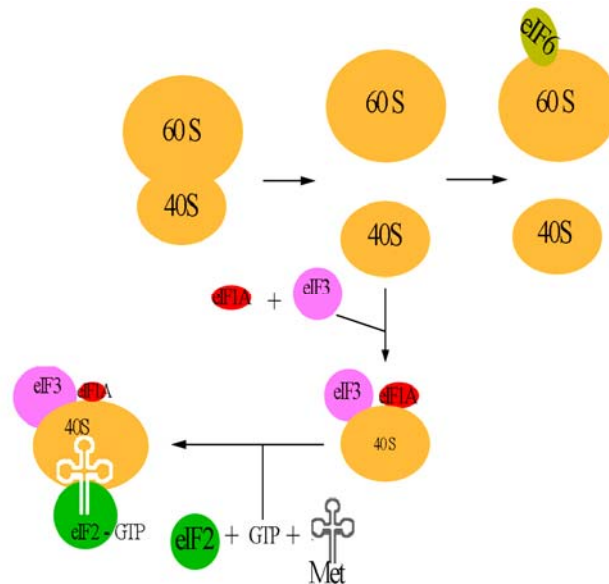
### 1.2.1. El modelo del rastreo del ribosoma (*scanning*).

El modelo del rastreo del ribosoma describe la forma en la que se inicia la traducción de la mayoría de los ARNm. En eucariotas se requieren al menos 11 factores diferentes para el inicio de la traducción. Todos ellos se encargan de que el metionil-ARN de transferencia iniciador (ARNt-Met) se sitúe sobre el triplete de iniciación que, generalmente, está constituido por la secuencia AUG. En los siguientes apartados se describe de forma abreviada este proceso (Sonenberg and Dever, 2003). El inicio de la traducción implica tres etapas:

**Formación del complejo de preiniciación de 43S.** Este complejo se forma gracias al factor eIF2 que permite la unión del ARNt-Met a la subunidad de 40S del ribosoma. Este proceso se detalla en la figura 5. Se sabe que el estado de fosforilación de este factor ante diferentes

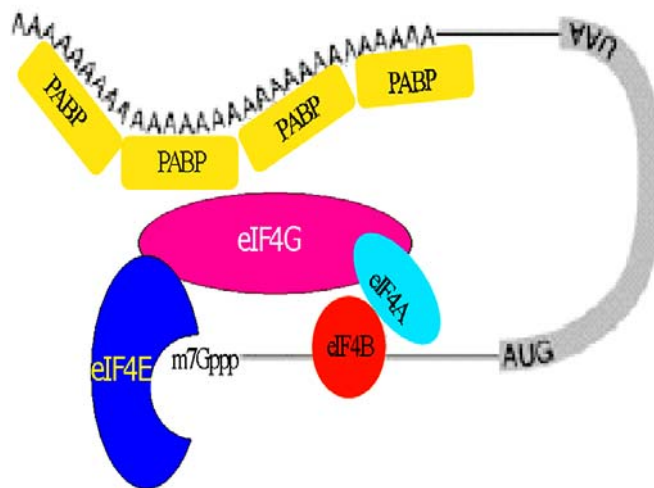


situaciones celulares, como pueden ser la carencia de factores de crecimiento, la infección por virus, el choque térmico o la entrada en división, hacen que disminuya la tasa de traducción celular.



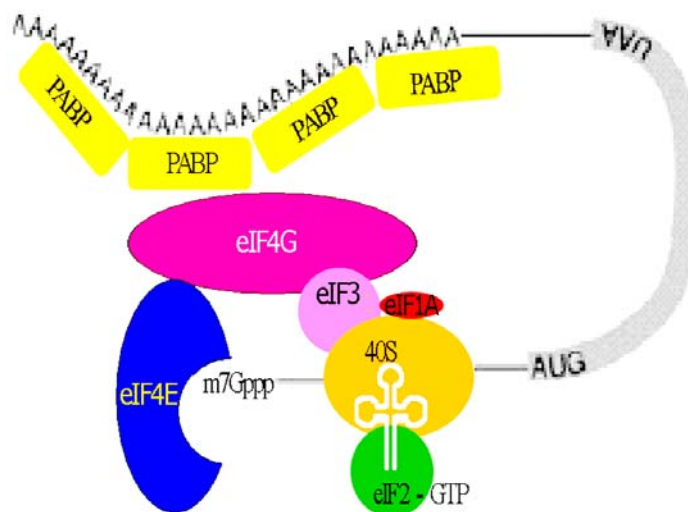
**Figura 5. Formación del complejo de preiniciación de 43S.** Dos factores, eIF3 y eIF6 mantienen separadas las subunidades de 40S y 60S del ribosoma. El complejo de preiniciación de 43S se forma cuando el complejo formado por eIF2, GTP y RNAt-Met se asocia con la subunidad 40S del ribosoma. Ésta a su vez se une a eIF3 y eIF1A, estabilizando este complejo. El factor eIF2 en la forma eIF2-GDP es inactivo; es necesaria su unión a GTP para que una el ARNt-Met.

**Unión del complejo de preiniciación al extremo 5' del ARNm.** Este proceso es llevado a cabo por los factores eIF3 y eIF4 (formado a su vez por eIF4A, eIF4E y eIF4G). En organismos eucariotas el extremo 5' se caracteriza por la existencia de una estructura denominada CAP (caperuza). Ésta está constituida por una metil guanosina trifosfato  $m^7Gppp$  que facilita el comienzo de la traducción y el transporte del ARN al citoplasma. Además de esta estructura CAP, la cola de poli-A de la que hemos hablado en el apartado 1.2.2.2 también estimula el proceso de la traducción cooperando con la estructura CAP (fig. 6). Uno de los factores proteicos implicados en este proceso, denominado eIF4G, funciona como un adaptador múltiple acercando varios componentes de la maquinaria necesaria para la traducción. El factor eIF4G es capaz de acercar los dos extremos de un ARNm a través de una interacción simultánea con la proteína de unión a la estructura CAP, denominada eIF4E, y con la proteína de unión a poli-A (PABP). Se forma de esta manera una estructura pseudo-circular necesaria para que tenga lugar la traducción. Se produce a continuación la unión del complejo de preiniciación 43S a esta estructura (fig. 7), mediante el consumo de ATP.



**Figura 6. Unión de eIF4F al inicio de la traducción.** Se muestra un ARNm con la estructura CAP (m7Gppp) y la cola de poli-A en los extremos 5' y 3', respectivamente. La unión del complejo eIF4F requiere la unión de eIF4E, eIF4G y eIF4A a la estructura CAP. eIF4A, estimulado por eIF4B, elimina la estructura secundaria de la zona 5' UTR del ARNm. La interacción de la proteína PABP con el factor eIF4G permite la circularización del ARNm.

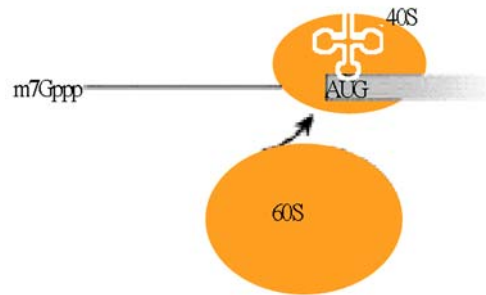
**Rastreo y reconocimiento del triplete iniciador AUG.** Es necesario que el complejo de preiniciación migre desde el extremo 5' del ARNm, en un proceso denominado rastreo, deslizándose por la zona 5' no traducida del ARNm (5'-UTR) hasta encontrar un triplete iniciador de la traducción AUG en un contexto idóneo. Así, el AUG iniciador es reconocido eficientemente como codón de iniciación cuando está incluido en la secuencia GCCPuCCAAUGG (Kozak, 2002).



**Figura 7. Unión del complejo 43S.** El complejo de preiniciación se une al ARNm mediante la interacción de eIF3 con eIF4G y con la subunidad ribosomal 40S unida a eIF2-GTP y al ARNt-Met. El complejo formado rastrea la zona 5' no traducida del ARNm hasta encontrar el codón de iniciación de la traducción. En ese momento se liberan los factores eIF2 y eIF3.

De esta secuencia (conocida como la secuencia de Kozak), la guanina en posición +4 (teniendo en cuenta que la adenina del codón de iniciación se numera como 1) es determinante para la selección del triplete iniciador así como la purina en posición -3, que es preferiblemente una adenina.

**Unión de la subunidad ribosomal de 60S.** Una vez que el complejo de preiniciación ha alcanzado el codón iniciador (fig. 8), el apareamiento de éste con el anticodón del ARNt-Met desencadena una serie de eventos que culminan con la unión de la subunidad ribosomal de 60S, formándose un ribosoma activo competente para la elongación. Este proceso conlleva la hidrólisis de GTP y la actuación de diferentes factores.

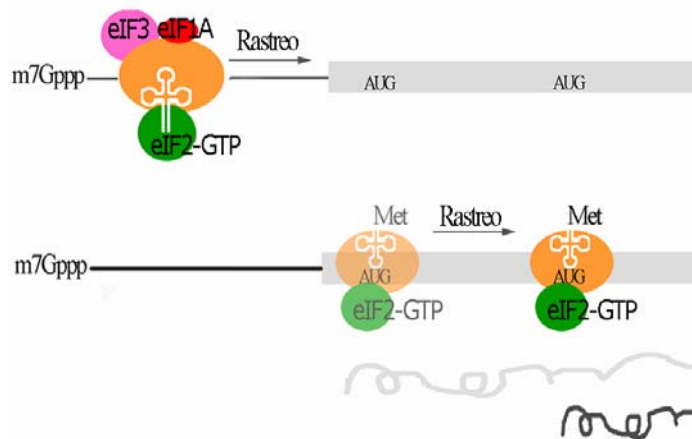


**Figura 8. Unión de la subunidad ribosomal de 60S.** El reconocimiento del anticodón por el Met-tARN desencadena la unión de la subunidad ribosomal de 60S.

A pesar de que este modelo es el tradicionalmente aceptado y mediante el cual tiene lugar la traducción de la mayor parte de los ARNm, existen mecanismos alternativos para el inicio de la traducción (Jackson, 2005). El análisis de las regiones 5' de diferentes ADN codificantes (ADNc) completos, muestra que el 49% de ellos presentan al menos un AUG en la secuencia anterior a la zona codificante. Esto significa que existen sitios potenciales para el inicio de la traducción en muchos ARNm, anteriores al utilizado comúnmente.

### 1.2.2. *Leaky scanning.*

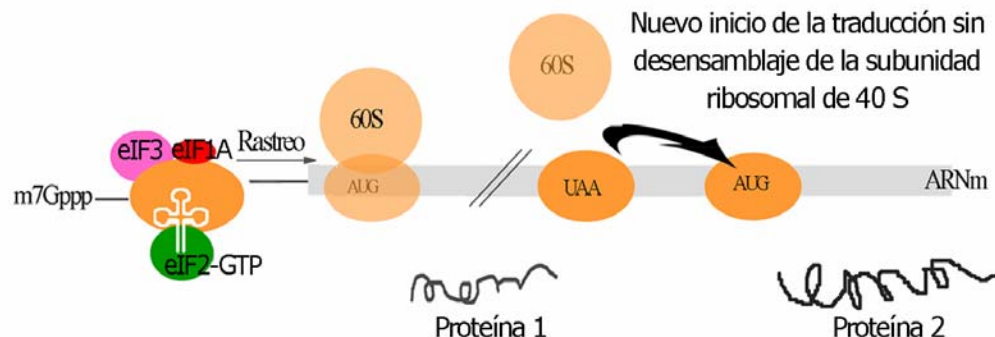
Algunos ARNm no comienzan la traducción en el primer AUG, sino en un triplete iniciador más interno. Esto sucede cuando el codón de iniciación está en un entorno débil para la traducción ya que la secuencia que lo rodea presenta diferencias respecto a la secuencia óptima de Kozak. Este proceso se conoce con el nombre de "*Leaky scanning*". Esta estrategia es utilizada por algunos virus pero está menos extendida entre eucariotas. Mediante este mecanismo algunos ARNm pueden producir dos proteínas que difieren en su extremo amino terminal o bien producir dos proteínas diferentes (fig. 9). Un ejemplo de este proceso ocurre en el ARNm que codifica para dos isoformas ( $\alpha$  y  $\beta$ ) del factor de transcripción C/EPB. En otros casos se sabe que la existencia de AUGs anteriores al utilizado como iniciador, regulan que el AUG desde el que comienza la traducción se utilice en menor medida y por lo tanto actúa limitando la cantidad de proteína que se sintetiza (Calkhoven et al., 2000).



**Figura 9. Leaky scanning.** El ribosoma rastrea la región 5' UTR del ARNm (panel superior). Al llegar al primer AUG (panel inferior), si éste no está en un contexto favorable, el ribosoma puede deslizarse por encima sin reconocerlo (en tenue) hasta otro AUG utilizado como codón de iniciación. Se pueden generar dos proteínas diferentes en función del AUG desde el cual comience la traducción.

### 1.2.3. Reiniciación.

Este proceso se basa en el inicio de la traducción desde un AUG interno, después de haberse traducido un fragmento de ARNm anterior (Jackson, 2005). No se conoce con exactitud el mecanismo mediante el cual sucede, pero puede generar dos proteínas diferentes a partir de un único ARNm (fig. 10).

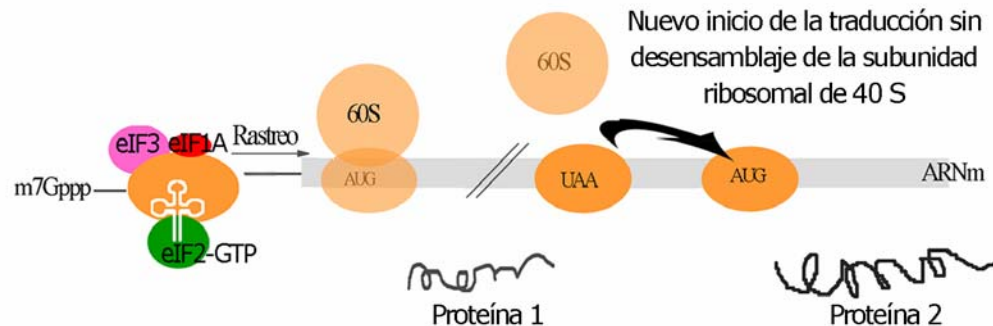


**Figura 10. Esquema de reiniciación.** El complejo de iniciación reconoce la estructura CAP (m7Gppp) y escanea la región 5'UTR hasta el codón de iniciación (AUG tenue). Se une la subunidad ribosomal de 60S y comienza la elongación. Al terminar la traducción, la maquinaria no se desensambla por completo y la subunidad ribosomal de 40S es capaz de empezar de nuevo la traducción en un codón de iniciación cercano.

La reiniciación es muy abundante en procariotas pero no en eucariotas. Diversos factores influyen en este proceso como la longitud de la secuencia anterior traducida y la distancia entre los dos sitios iniciadores. El ARNm que codifica para el factor 4 activador de transcripción (ATF4) de metazoos constituye un ejemplo de reiniciación (Vattem and Wek, 2004).

#### 1.2.4. *Shunting* (salto del ribosoma).

El *shunting* consiste en un mecanismo de inicio de la traducción en el que el ribosoma se une al ARNm reconociendo la estructura CAP pero “salta” una zona más o menos larga de la región 5'-UTR, ignorando los posibles tripletes de iniciación que ésta contenga. Las zonas en las que sucede esto suelen estar altamente estructuradas. Un esquema de este proceso se muestra en la figura 11. Este mecanismo fue descrito en la traducción del virus del mosaico de la coliflor (Futterer et al., 1993). Se han descrito también ejemplos de este mecanismo en ARNm de mamíferos como es el caso de  $\beta$ -secretasa que posee tres AUG anteriores al utilizado para la síntesis de esta proteína.



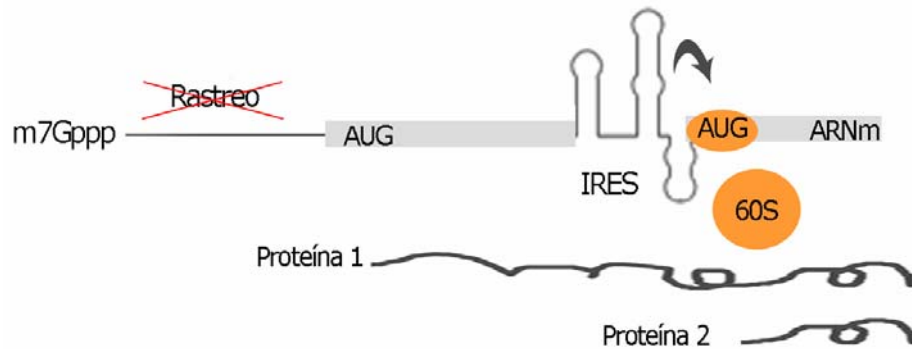
**Figura 11. Esquema del *Shunting*.** Se representa un ARNm en el que el complejo de iniciación reconoce la estructura CAP (m7Gppp) y comienza el rastreo de la zona 5' UTR. La existencia de una estructura secundaria compleja, hace que el ribosoma no detecte el primer codón de iniciación (en tenue) y se dirija al siguiente sin rastrear la zona con alta estructura secundaria. En este segundo AUG comienza la traducción.

#### 1.2.5. Sitios de entrada interna del ribosoma (IRES).

Se conoce otro tipo de iniciación de la traducción, diferente del modelo de rastreo, que se caracteriza por no necesitar el reconocimiento de la estructura CAP del extremo 5' del ARNm. Existe una secuencia comprendida en el ARNm a la que se ancla directamente el ribosoma, sin necesidad de rastrear la zona 5' anterior a dicha secuencia (fig. 12). Esta zona del ARNm capaz de reclutar de forma directa el ribosoma se conoce como sitio de entrada interna del ribosoma, IRES, (*Internal Ribosome Entry Site*). Este tipo de traducción se describió inicialmente en picornavirus aunque en la actualidad son muchos los ejemplos de ARNm eucariotas que utilizan este tipo de traducción, como por ejemplo el ARNm que codifica para la proteína PITSLRE, una proteína quinasa perteneciente a la familia de quinasas dependientes de ciclina (Cornelis et al., 2000).

La utilización de este tipo de traducción en eucariotas ha sido descrita en procesos en los que la traducción depende de la estructura de

CAP está impedida. Esto sucede durante apoptosis, división o en aquellas condiciones que suponen un estrés celular.



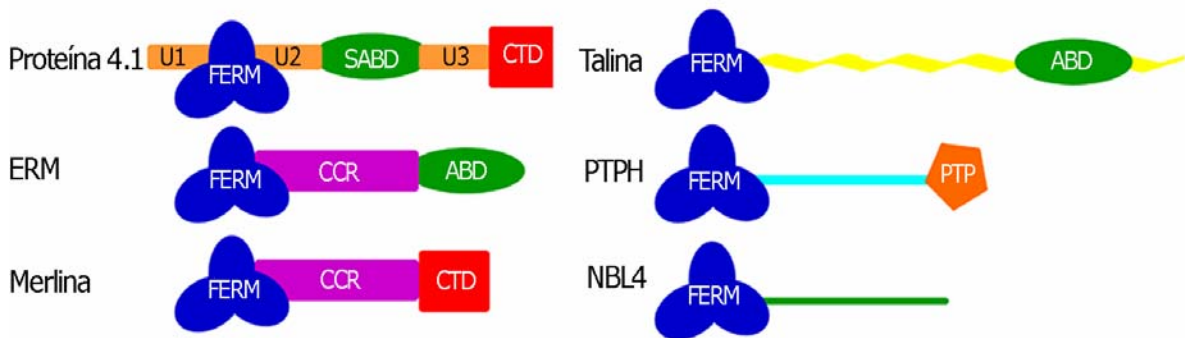
**Figura 12. Sitio de entrada interna para el ribosoma (IRES).** Se muestra un esquema de un ARNm en el que la subunidad ribosomal 40S es reclutada directamente por una secuencia del propio ARNm denominada IRES. Se sabe que son necesarios algunos de los factores canónicos del inicio de la traducción (no se muestra). En este tipo de iniciación no es necesario el reconocimiento de la estructura CAP (m7Gppp) ni el rastreo de la zona 5'UTR. Algunos ARNm pueden codificar de esta forma para dos proteínas que, dependiendo de si los AUG iniciadores están o no en fase, serán dos proteínas diferentes o isoformas de una misma proteína con distinto extremo amino terminal.

Se conocen en la actualidad algunas proteínas de unión al ARN que regulan la actividad IRES concretos. Estas proteínas se denominan factores transactivadores (ITAFs, *trans-acting factors*). Algunos ejemplos de estos ITAFs son las proteínas PTB (*polypyrimidine tract binding protein*) (Giraud et al., 2001); el autoantígeno La (Holcik and Korneluk, 2000) y hnRNP A1 (Bonnal et al., 2005). Algunos IRES poseen en su región 3' una región rica en polipirimidinas importante para su actividad. Por ejemplo el IRES presente en el virus de la hepatitis C (HCV) presenta la secuencia CYYYYCYYYY(G/Y)G, donde Y representa una pirimidina, importante para la actividad de este IRES viral. Los ITAFs parecen ser ubicuos y aún no se ha establecido como actúan. No son los únicos que afecta a la actividad de los IRES, sino que los factores canónicos de la traducción también lo hacen. Así, por ejemplo, cuando el factor de traducción eIF4E se elimina, se facilita el uso del IRES del ARNm que codifica para la proteína *Reaper* de *Drosophila*. A pesar de que en la actualidad no se conocen secuencias conservadas que puedan predecir la existencia de un IRES en un ARNm, sí se conocen algunos dominios que se mantienen en IRES concretos como la zona de pirimidinas que hemos comentado anteriormente. Por lo general los IRES se ubican en regiones con una gran estructura secundaria.



## 2. Proteína 4.1R: un miembro de la superfamilia 4.1.

La proteína 4.1R fue la primera descrita de una superfamilia de proteínas a la que dio nombre: la superfamilia de 4.1. Todas las proteínas que componen esta superfamilia se caracterizan por poseer un dominio muy conservado en la zona amino denominado FERM (*Four point one protein, Ezrin, Radixin, Moesin*) a través del cual interaccionan con glicoproteínas de la membrana plasmática (Chishti et al., 1998). Sus miembros se agrupan en cinco familias; la familia de las proteínas 4.1, la de las ERM (*Ezrin, Radixin, Moesin*); la talina, las proteínas tirosina fosfatasas (PTPH) y las proteínas NBL4 (*Novel Band 4.1-Like 4*) (fig. 13) (Calinisan et al., 2006). Algunas de estas familias, como las ERM y la talina, además de contener el dominio FERM, mantienen otro dominio muy conservado denominado ABD que permite la interacción con la actina. En la proteína 4.1R este dominio se denomina SABD ya que permite la interacción con la actina y con la espectrina.



**Figura 13. Esquema de los dominios estructurales de algunos miembros de la superfamilia de 4.1.** Todos sus miembros se caracterizan por poseer un dominio amino con alta homología, FERM. El dominio SABD o ABD permiten la unión con actina o con actina y espectrina respectivamente, y lo presentan la proteína 4.1 las ERM y la talina. Las ERM y la merlina presentan también un dominio conservado CCR (*Coiled-Coiled Region*). CTD representa el dominio carboxilo terminal, PTP dominio fosfatasa y U1, U2 y U3 regiones únicas.

La familia de las proteínas 4.1 está compuesta a su vez por cuatro proteínas diferentes denominadas 4.1R, 4.1B (DAL), 4.1N y 4.1G codificadas por genes diferentes y con un patrón de expresión diferencial (tabla 1) (Hoover and Bryant, 2000). Todas ellas se caracterizan por poseer, además del dominio FERM y el dominio SABD de unión a actina y espectrina, un dominio conservado denominado CTD en la zona carboxilo terminal y tres regiones variables denominadas U1, U2 y U3. Todas las proteínas de esta familia presentan multitud de isoformas diferentes debido principalmente al procesamiento alternativo de su pre-ARNm. Es posible que las regiones variables U1, U2 y U3 confieran la especificidad de determinadas funciones e interacciones en las diferentes isoformas de las proteínas 4.1.

Nombres de la proteína	Peso Molecular	Cromosoma	Distribución en tejido adulto	Localización subcelular
4.1R, EPB41	80, 135	1p36.2-p34	Eritrocitos, Cerebro (Shi et al., 1999) ; (Walensky et al., 1998)	Membrana plasmática, núcleo, polos del huso mitótico, centrosomas
4.1B, DAL-1, EPB41L3	125, 145	18p11.3	Cerebro, corazón, pulmón, riñones, intestino, testículos y glándula adrenal (Parra et al., 2000) ; (Peters et al., 1998)	Membrana plasmática y en contactos célula-célula
4.1N, EPB41L1	100, 135	20q11.2-q12	Cerebro y nervios periféricos (Walensky et al., 1999; Yamakawa et al., 1999)	En el aparato mitótico, núcleo, en membrana, en dendritas neuronales
4.1G, EPB41L2	95, 140	6q23	Corazón, placenta, cerebro, pulmón, músculo esquelético, riñón, páncreas y gónadas. (Parra et al., 1998; Walensky et al., 1998)	Membrana plasmática

**Tabla 1. Familia de las proteínas 4.1.** Se indican los nombres de las cuatro proteínas que forman la familia de las 4.1 (4.1R, 4.1B, 4.1N y 4.1G), así como sus pesos moleculares, los cromosomas en los que se encuentran los genes que codifican para estas proteínas y la distribución que presentan las isoformas descritas hasta el momento en tejidos y líneas celulares.

## 2.1. La proteína 4.1R.

La proteína 4.1 fue aislada inicialmente de eritrocitos. El análisis de los fantasmas de eritrocito mediante SDS-PAGE daba lugar a una serie de bandas proteicas entre las que se encontraba 4.1R (Fairbanks et al., 1971); (Steck, 1974). El nombre que recibe esta proteína se debe a la posición que ocupaba en estos geles, donde aparecía como un doblete con pesos moleculares aparentes de 78 y 80 kDa (Conboy, 1993). En la actualidad se sabe 4.1R contribuye a mantener la forma de esta célula y las propiedades mecánicas de su membrana plasmática.

Estudios en células nucleadas permitieron establecer que 4.1R no se encontraba únicamente en eritrocitos sino también en células nucleadas de metazoos (Anderson et al., 1988) (Aster et al., 1984; Cohen et al., 1982; Davies and Cohen, 1985; Goodman et al., 1984; Granger and Lazarides, 1984; Granger and Lazarides, 1985; Leto et al., 1986; Spiegel et al., 1984; Tang et al., 1988). En estas células 4.1R presentaba diferentes isoformas cuyos pesos moleculares aparentes oscilaban desde los 30 hasta los 210 kDa. Se ha determinado la presencia de diferentes isoformas de 4.1R en diversos tejidos



(nervioso, digestivo, sensorial, reproductor, muscular, entre otros) y en distintas especies (insectos, anfibios, aves y mamíferos). Todas estas isoformas se han subdividido en tres grupos en función del lugar desde el cual tiene lugar el comienzo de la traducción en los ARNm que las generan (ver más adelante). Las isoformas de peso molecular alto con un peso molecular aproximado de 135 kDa, las de peso molecular intermedio, de aproximadamente 80 kDa y el grupo de isoformas de peso molecular bajo con un peso molecular aproximado de 60 kDa (Luque and Correas, 2000).

## 2.2. *EPB41*: el gen de 4.1R

El gen que codifica para la proteína 4.1R se clonó en 1977, abarca unos 200 kpb y se encuentra localizado en el cromosoma 1 (1p33-p34.2). Inicialmente se localizaron 23 exones individuales y su región 3' no traducida, de unos 3kpb (Baklouti et al., 1997). En el año 1997 se identificó otro exón, el E17B (Schischmanoff et al., 1997) y más recientemente se han identificado otros 3 exones en la zona 5' de este gen, denominados 1A, 1B y 1C (Parra et al., 2003). En la parte superior de la figura 14 se detalla la composición exónica de este gen.

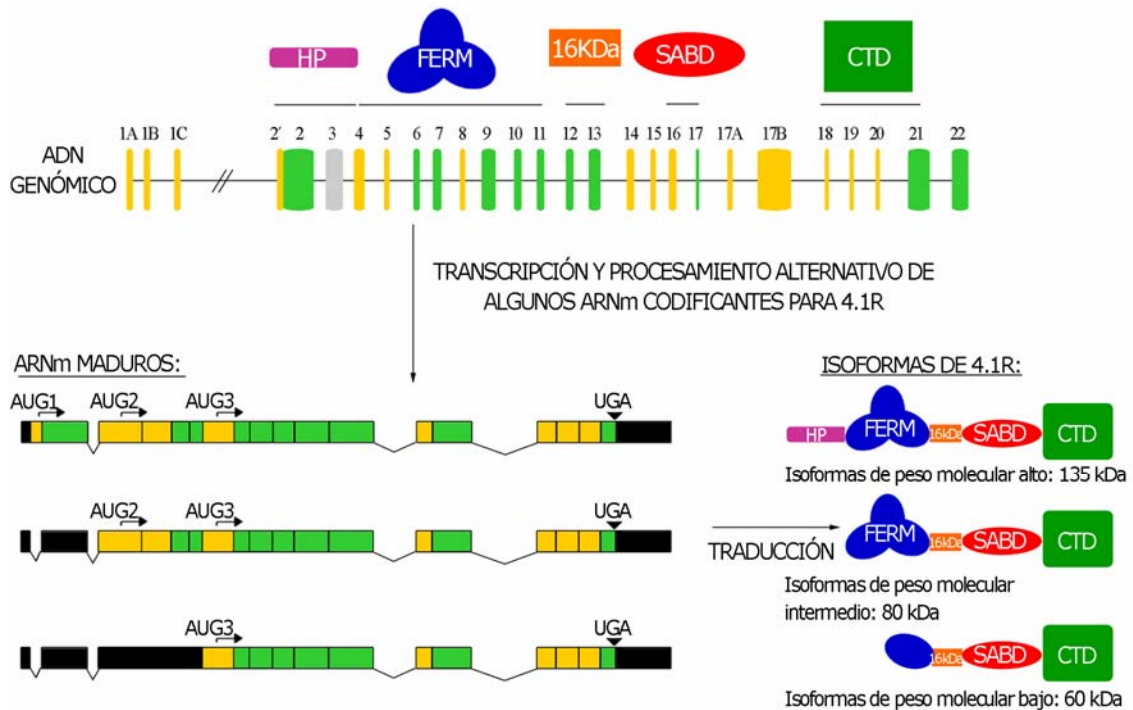
## 2.3. Procesamiento alternativo del pre-ARNm de 4.1R.

El gran número de isoformas de 4.1R se genera a partir del procesamiento alternativo que sufre su pre-ARNm. De sus 27 exones, 15 son exones alternativos (1A, 1B, 1C, 2', 4, 5, 8, 14, 15, 16, 17A, 17B, 18, 19 y 20), lo que confiere al pre-ARNm de 4.1R un gran potencial para generar isoformas diferentes de 4.1R. Como ya se ha comentado, en el eritrocito existe una isoforma de 4.1R de 80 kDa pero las células nucleadas pueden presentar un amplio número de isoformas de 4.1R diferentes dependiendo del tejido y del momento de desarrollo en que se presenten. A continuación se detalla el procesamiento de algunos de sus exones de forma más concreta:

**Procesamiento del exón 2' y los exones 1A, 1B y 1C.** Este procesamiento consiste en la inclusión o exclusión de 17 nucleótidos situados en la región 5' del exón 2. Estos 17 nucleótidos se han denominado exón 2' (Conboy et al., 1991a). Todos aquellos ARNm maduros que contienen este exón son traducidos a partir de un codón de iniciación para la traducción presente en él, denominado ATG1. Se generan así las isoformas de 4.1R que hemos denominado de peso molecular alto (aproximadamente 135 kDa) (fig. 14). La exclusión de este exón genera ARNm maduros que se traducen desde un triplete de iniciación situado en el exón 4, denominado ATG2. Las isoformas generadas desde este codón de inicio de la traducción constituyen el grupo de 4.1R englobadas en isoformas de peso molecular intermedio (aproximadamente unos 80 kDa).

Estudios recientes han determinado la existencia de tres exones excluyentes entre sí (1A, 1B y 1C) localizados en la región 5' de los ARNm maduros de 4.1R (Parra et al., 2003). Estos primeros exones excluyentes no se traducen

pero su elección parece determinar la presencia del exón 2' en los ARNm maduros de 4.1R. Así, todos aquellos ARNm aislados hasta la fecha que poseen el exón 1A carecen del exón 2', mientras que aquellos ARNm que tienen como primer exón el 1B o 1C poseen siempre el exón 2'.



**Figura 14. Gen, transcritos e isoformas de 4.1R.** A) Se ha representado el gen que codifica para la proteína 4.1R. Los exones constitutivos se han representado en verde y los alternativos en amarillo. Encima de la composición exónica indican los exones concretos que codifican para los diferentes dominios de 4.1R (HP, FERM, dominio de 16 kDa, SABD y CTD). Tras la transcripción de este gen se produce un pre-ARNm (no se muestra) que está sujeto a un amplio procesamiento alternativo. B) A pesar de que son muchos los ARNm maduros a los que este pre-ARN puede dar lugar, se han clasificado en tres tipos en función del codón que utilizan para el inicio de la traducción. Aquellos que mantienen el exón 2' y por lo tanto contienen el codón de iniciación AUG1; los que carecen de este exón 2' pero mantienen el exón 4 que contiene el codón de iniciación AUG2 y aquellos que pierden el exón 2' y 4 y utilizan para el inicio de la traducción un codón iniciador presente en el exón 8, AUG3. Cada uno de estos grupos de ARNm da lugar a una isoforma concreta de 4.1R que puede ser clasificada como isoforma de peso molecular alto, intermedio o bajo (135, 80 y 60 kDa, respectivamente) en función del codón de iniciación desde el cual se ha traducido el ARNm del que provienen.

En cuanto a la importancia de este procesamiento hay que señalar que los ARNm maduros codificantes para 4.1R eritroides eliminan el exón 2', generando isoformas de 80 kDa (Chasis et al., 1993). En cambio, en células no eritroides son más abundantes las isoformas de 4.1R de peso molecular alto (aproximadamente 135 kDa).

**Procesamiento del exón 16.** La inclusión o exclusión del exón 16, que codifica para una región de 21 aminoácidos localizada en el dominio de 10

kDa o SABD, da lugar a un dominio de unión a actina y espectrina completo o incompleto, respectivamente (Baklouti et al., 1996; Conboy et al., 1991a; Tang et al., 1990). La inclusión de este exón por procesamiento alternativo se ve favorecida en células eritroides, produciéndose además en estadios tardíos de la eritropoyesis. La expresión de este exón es esencial para la correcta localización de 4.1R en el citoesqueleto submembranal del eritrocito y para la formación del complejo actina-espectrina-4.1R.

## 2.4. Dominios estructurales de la proteína 4.1R.

Se han definido diferentes dominios en la proteína 4.1R de eritrocito mediante su digestión parcial con quimiotripsina. Los dominios son los siguientes:

**El dominio HP.** Puede llegar a tener 209 aminoácidos y no está presente en la proteína eritroide de 80 kDa. Las isoformas de 4.1R que portan este dominio se sintetizan a partir del codón de iniciación de la traducción (ATG1) situado en el exón 2', sujeto a procesamiento alternativo (Tang et al 1990). Mediante este dominio 4.1R se une a la proteína centrosomal CPAP, (*Centrosomal P4.1 Associated Protein*) y se ha sugerido que 4.1R podría estar anclando el complejo CPAP/ $\gamma$ -tubulina al centrosoma de la célula. Se ha descrito también en este dominio un posible sitio de unión a calmodulina identificado por homología con la proteína 4.1 de *Xenopus* (Kelly et al., 1991) cuya funcionalidad ha sido objeto de estudios *in vitro* (Kelly et al., 1991; Leclerc and Vetter, 1998; Vetter and Leclerc, 2001).

**El dominio FERM.** La cristalización de este dominio de 30 kDa ha permitido establecer que adopta una conformación trilobular (Han et al., 2000). El dominio amino del FERM, denominado "dominio N", tiene un sitio de unión con la banda 3 que permite la unión de 4.1R a la membrana plasmática en el eritrocito (Jons and Drenckhahn, 1992; Pasternack et al., 1985). En el lóbulo del dominio FERM, denominado "dominio alfa", se encuentra el sitio de unión con la glicoforina A y C (Anderson and Lovrien, 1984; Marfatia et al., 1995; Marfatia et al., 1994). El lóbulo carboxilo del dominio FERM, denominado dominio C, contiene el sitio de unión a miembros de la familia MAGUK, proteínas guanilato quinasas asociadas a la membrana. Éstas reclutan y organizan a otras proteínas de membrana, coordinando vías de señales de transducción, entre las que se encuentran CASK (Marfatia et al., 1995) y hDlg (Lue et al., 1994). El "dominio C" del FERM contiene además un sitio para unión a tubulina (Perez-Ferreiro et al., 2001) que permite a la proteína 4.1R asociarse a los microtúbulos interfásicos.

La proteína 4.1R puede también asociarse a la calmodulina mediante el dominio FERM (Tanaka et al., 1991), permitiendo una regulación mediada por la presencia del catión  $\text{Ca}^{2+}$ . La unión de  $\text{Ca}^{2+}$  a la calmodulina altera la conformación de los tres lóbulos del FERM y debilita la unión con p55, glicoforina C (GPC) y Banda 3.

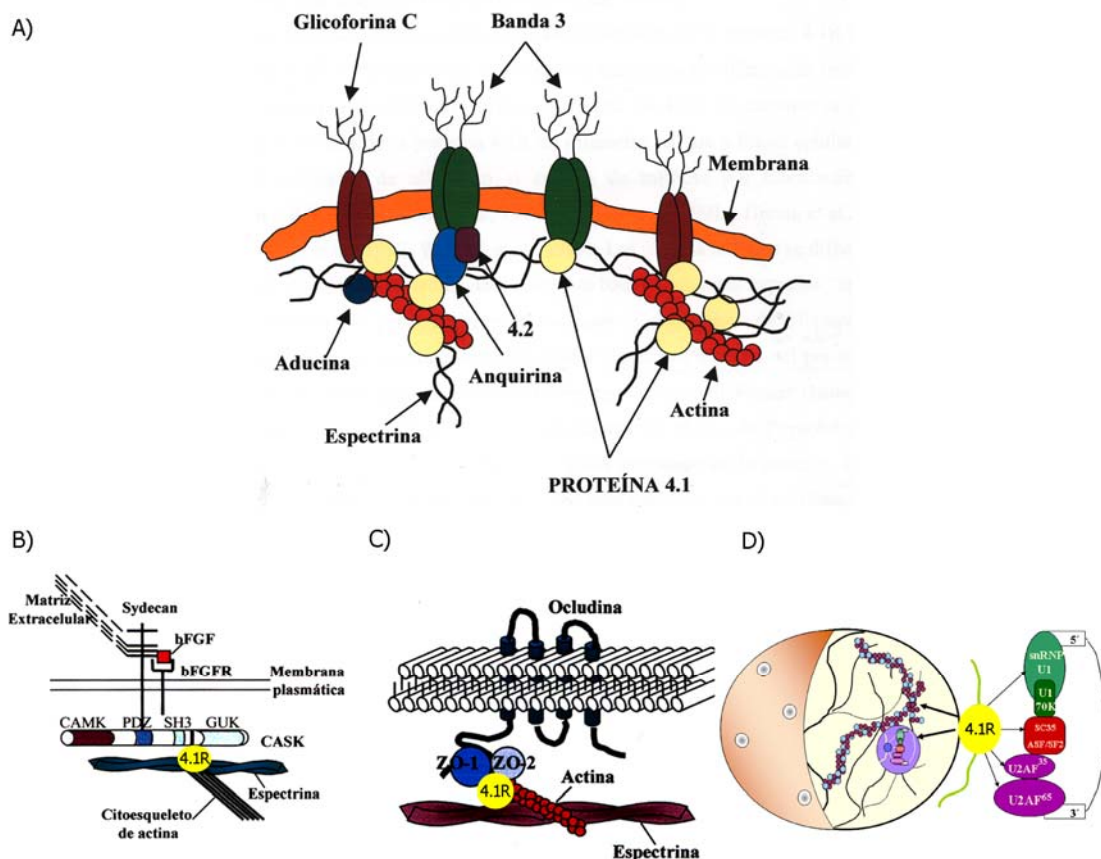
**El dominio de 16 kDa.** No se han descrito interacciones con otras proteínas a través de este dominio pero se sabe que contiene sitios de fosforilación específicos que pueden modular la afinidad de regiones adyacentes de 4.1R con sus ligandos (Danilov et al., 1990; Eder et al., 1986; Ling et al., 1988). Así, la fosforilación de la serina 312 de este dominio por la proteína quinasa C (PKC), disminuye la unión del dominio FERM con las proteínas p55, GPC y banda 3. La fosforilación de esta serina induce también un cambio conformacional en el dominio SABD (ver más adelante) que disminuye la interacción de 4.1R con la membrana plasmática a través de la actina.

**El dominio SABD (*Spectrin Actin Binding Domain*) o de 10 kDa.** Este dominio posee un sitio de unión a la actina y la espectrina (Correas et al., 1986a; Correas et al., 1986b), además de una tirosina fosforilable por proteínas quinasas capaz de modular dicha interacción (Subrahmanyam et al., 1991). También tiene sitios de unión a la miosina (Pasternack and Racusen, 1989) y a la tropomiosina, interacción importante para el mantenimiento del aparato contráctil de las células del músculo estriado (Kontogianni-Konstantopoulos et al., 2001a).

**El dominio CTD (*Carboxy Terminal Domain*).** Sólo se conocen interacciones de este dominio en células nucleadas, donde la proteína 4.1R interacciona con la proteína NuMa, componente de la matriz nuclear y del aparato mitótico (Mattagajasingh et al., 1999). A través de este dominio se ha descrito también la interacción con el factor eIF3-p44 (Hou et al., 2000), imprescindible para el inicio de la traducción. Otra proteína que interacciona con 4.1R a través de este dominio es la proteína ZO-2 (Mattagajasingh et al., 2000), anclando el citoesqueleto de actina con las uniones estrechas.

## 2.5. Funciones de la proteína 4.1R en el eritrocito.

En los eritrocitos la isoforma de 80 kDa de 4.1R contribuye a mantener la forma de esta célula y las propiedades mecánicas de su membrana plasmática. Esta membrana requiere una alta elasticidad y flexibilidad que permita al eritrocito viajar por los finos capilares para transportar el oxígeno a todas las regiones del cuerpo. Esto se consigue gracias a que



**Figura 15. Funciones de la proteína 4.1R.** A) En el eritrocito 4.1R ancla el citoesqueleto submembranar con proteínas integrales de membrana. La proteína 4.1R interacciona con los filamentos de actina, con la espectrina y con el dominio citoplasmático de las proteínas integrales de membrana banda 3 y glicoforina C. La anquirina interacciona con la espectrina y con el dominio citoplasmático de la banda 3. Se representan otros componentes del citoesqueleto submembranar como la aducina y la proteína 4.2. B) Asociación de la matriz extracelular y el citoesqueleto en células no eritroides. La proteína 4.1R media la asociación entre la matriz extracelular y el citoesqueleto de actina a través de su interacción con la proteína humana CASK, componente de la familia MAGUK. La proteína CASK interacciona con el dominio FERM de la proteína 4.1R a través de una región situada entre su dominio SH3 y GUK (Cohen et al 1998). C) La proteína 4.1R actúa como nexo de unión entre las proteínas de las uniones estrechas y el citoesqueleto de actina en células epiteliales. La proteína 4.1R se une a través de su dominio CTD a la proteína ZO-2 (Mattagajasingh et al 2000). D) La proteína 4.1R es un componente del nucleoesqueleto al que se asocian componentes de la maquinaria del procesamiento del pre-ARNm en células nucleadas. La proteína 4.1R se encuentra en la matriz nuclear (De Cárcer et al 1995). Además se han descrito interacciones con snRNP, el factor SC35 y el factor U2AF (Lallena et al., 1998).



el citoesqueleto submembranal (Bennett and Gilligan, 1993) (denso esqueleto fibroso formado por los dominios de las proteínas situadas en la zona citoplasmática de la membrana plasmática como la espectrina y la actina) interacciona con proteínas como 4.1R. Ésta a su vez se une a proteínas integrales de membrana como son la glicoforina C y la banda 3 (fig. 15A).

Así, 4.1R sirve de nexo de unión entre el citoesqueleto submembranal y la membrana plasmática. Alteraciones en los niveles de 4.1R o en las interacciones con otras proteínas pueden conllevar el desarrollo de eliptocitosis hereditaria (Conboy et al., 1991); (Lambert et al., 1988). Esta enfermedad se manifiesta por la aparición de eritrocitos con forma elíptica con menor capacidad de deformación y estabilidad, que acaban fragmentándose con el consiguiente desarrollo de anemia hemolítica (Shi et al., 1999).

## 2.6. Función de la proteína 4.1R en células nucleadas.

A partir de las diferentes localizaciones intracelulares y de las interacciones proteína-proteína en las que participa 4.1R en células no eritroides se han caracterizado algunas de sus posibles funciones.

**Membrana plasmática.** Parece que el papel que 4.1R desarrolla en el mantenimiento de las propiedades de la membrana plasmática no es exclusivo de las células eritroides sino que se mantiene también en las células no eritroides. La proteína 4.1R interacciona mediante su dominio FERM con proteínas de la familia MAGUK (proteínas guanilato quinasas asociadas a la membrana plasmática) entre las que se encuentran hDlg (proteína humana homologa a Dlg, *Disc large gene de Drosophila melanogaster*), la proteína CASK (Cohen et al., 1998) que a su vez interaccionan con la membrana plasmática. La doble unión de 4.1R a estas proteínas y a la actina la hacen funcionar como nexo entre el citoesqueleto celular y la membrana plasmática (fig. 15B).

**Uniones estrechas.** A través del dominio CTD 4.1R interacciona con las proteínas MAGUK de las uniones estrechas, ZO-1 y ZO-2. La unión de 4.1R a estas proteínas así como a ocludina, espectrina y actina sugieren su importancia en uniones estrechas y en la organización del citoesqueleto de células confluentes (fig. 15C).

**Músculo estriado.** La proteína 4.1R interacciona con la miosina, la  $\alpha$ -actina y la tropomiosina, todas ellas componentes principales del sarcómero del músculo estriado. Todas estas asociaciones indican que 4.1R está implicada en la organización estructural y el mantenimiento del aparato contráctil (Kontogianni-Konstantopoulos et al., 2000).

**Maquinaria de traducción.** La interacción de 4.1R con el factor eucariota eIF3-p44, esencial para el inicio de la traducción, sugiere que 4.1R

puede intervenir anclando la maquinaria de traducción al citoesqueleto (Hou et al., 2000).

**División celular.** Las interacciones que 4.1R mantiene con las proteínas NuMa (Mattagajasingh et al., 1999) y con la proteína centrosomal (CPAP) parecen implicar a 4.1R en el proceso de división celular.

**Matriz nuclear.** La proteína 4.1R forma parte de la matriz nuclear (De Cárcer et al., 1995), localizándose en dominios enriquecidos en factores implicados en el procesamiento alternativo del pre-ARNm como SC35. Existen evidencias de que 4.1R interacciona con factores de maduración del pre-ARNm como por ejemplo U2AF<sup>35</sup>, proteínas SM (componentes de las sRNPs) y proteínas SR. Todo esto implica a 4.1R en el mantenimiento de la arquitectura del núcleo y en la regulación de la actividad transcripcional.

Además de todo esto, se ha descrito la capacidad de algunas isoformas de 4.1R para autoasociarse, aunque la funcionalidad de esta asociación es todavía desconocida (Pérez-Ferreiro et al., 2006).

Uno de los mayores retos en la investigación de 4.1R es determinar los mecanismos generadores de diversidad de esta proteína y las funciones que desempeñan cada una de sus isoformas.

# Objetivos



## OBJETIVOS

El pre-ARNm de la proteína 4.1R tiene un gran potencial para formar diferentes ARNm maduros mediante procesamiento alternativo, generando una gran variabilidad de isoformas de esta proteína. En los últimos años nuestro grupo se ha centrado en el aislamiento de ADN codificantes para isoformas de 4.1R, contribuyendo a establecer los patrones de procesamiento del pre-ARNm de esta proteína. Los objetivos de este trabajo son:

1. Estudiar la posible implicación de un mecanismo de traducción interna en la generación de isoformas de 4.1R.
2. Aislar y caracterizar ADN codificantes para isoformas de 4.1R que carezcan del dominio CTD.

## Materiales y métodos

## **MATERIALES Y MÉTODOS**

### **1. Cultivos celulares.**

#### **1.1. Líneas celulares de mamífero.**

**COS-7.** Células fibroblásticas de riñón de mono transformado con el virus SV40. Estas células han sido usadas para analizar la expresión transitoria de diversos ADNc y para el estudio de proteínas 4.1R endógenas.

**MDCK.** Células epiteliales de riñón de perro. Empleadas para la transfección transitoria de diferentes ADNc de 4.1R, en el estudio de isoformas endógenas de esta proteína y para el aislamiento de ARN.

**HeLa.** Células epiteliales de cuello de útero humano. Empleadas para transfectar distintos ADNc y para el estudio de isoformas endógenas de 4.1R.

**HepG2.** Células epiteliales de hígado humano. Se han usado para transfectar distintos ADNc y para el estudio de isoformas endógenas de 4.1R.

**Jurkat.** Línea celular linfoblastoide T humana. Utilizadas para el estudio de isoformas endógenas de 4.1R.

**Raji.** Línea celular linfoblastoide B humana. Utilizadas para el estudio de isoformas endógenas de 4.1R.

**MOLT-4.** Células linfoides T humanas. Utilizadas en el estudio de isoformas endógenas de 4.1R sin dominio CTD.

**ECV304** Células de cordón umbilical humano. Empleadas para el estudio de isoformas endógenas de 4.1R.

Las células COS-7, HeLa, MDCK y HepG2 fueron cultivadas en medio Eagle modificado por Dulbecco (DMEM, Life Technologies, Inc.). Los cultivos de las líneas celulares no adherentes (MOLT-4, Jurkat y Raji) se realizaron en medio RPMI-1640 (Life Technologies, Inc.). Ambos medios fueron suplementados con suero fetal de ternera (Sigma-Aldrich) al 10% (v/v) en el caso de las líneas linfoides y al 5% en el caso de las células adherentes, glutamina al 1% (p/v), penicilina 50U/ml y estreptomicina 50U/ml. En el caso de células de hepatoma humano HepG2 además, se suplementó el medio con 4,5 g/l de glucosa. Los cultivos se mantuvieron a 37° C, en una atmósfera con el 95% de humedad relativa y enriquecida hasta un 5% (v/v) con CO<sub>2</sub>. Para despegar las células adherentes de la superficie de crecimiento, se utilizó una solución con tripsina al 0,05% y EDTA al 0,02%.

## 1.2. Cepas bacterianas.

Todas las cepas bacterianas usadas para el desarrollo de esta tesis son de *Escherichia coli*:

- DH5 $\alpha$  (Hanahan, 1983).
- TOP10 (Invitrogen).
- Epicurean Coli® XL1-Blue (Stratagene).
- BL21 (Amersham Pharmacia Biotech).

Los cultivos bacterianos se llevaron a cabo en los medios LB o SOC (Sambrook et al., 1989) o NZY+Broth; NZ amina (caseína hidrolizada) al 1% (p/v), extracto de levadura (DIFCO) al 0,5% (p/v), NaCl al 0,5% (p/v), 12,5 mM de MgCl<sub>2</sub>, 12,5 mM de MgSO<sub>4</sub> y 20 mM de glucosa. Los cultivos en medio líquido se crecieron con agitación a 225 rpm y 37° C. Los cultivos en medio sólido se realizaron sobre placas Petri (Bibby Sterilin), preparadas con medio LB y Agar Bacteriológico "Europeo" (Hispanlab) al 2% (p/v) y a 37° C. Cuando fue necesario los medios se suplementaron con ampicilina o kanamicina hasta una concentración final de 100 o 50  $\mu$ g/ml respectivamente.

## 2. Anticuerpos.

### 2.1. Anticuerpos primarios.

En el desarrollo de esta tesis se han utilizado los siguientes anticuerpos anti-4.1R:

**10b.** Anticuerpo policlonal purificado por afinidad. Obtenido por inmunización de conejos con el péptido sintético que comprende la secuencia codificada por el exón constitutivo 17, MESVPEPRPSEWDK, (Correas et al., 1986b).

**762.** Anticuerpo policlonal obtenido por inmunización de conejos con un péptido sintético que comprende una secuencia codificada por el exón constitutivo 2, TKNKERTSESRGC, sólo presente en las isoformas de 4.1R de peso molecular alto (Luque et al., 1999). La inmunización se llevó a cabo tal como se ha descrito con anterioridad (Correas et al., 1986b).

**10a.** Anticuerpo policlonal purificado por afinidad, obtenido por inmunización de conejos con un péptido sintético que comprende una secuencia codificada por el exón alternativo 16, KKRERLDGENIYIRHSNLMLE (Correas et al., 1986a).

**937.** Anticuerpo policlonal obtenido por inmunización de conejos con la zona comprendida entre el exón 17 y 22 de 4.1R.

Otros anticuerpos usados en el desarrollo de esta tesis se describen a continuación:

**9E10.** Anticuerpo monoclonal anti-c-myc, obtenido a partir de un hibridoma derivado de ratón, de la *American Type Culture Collection*. Reconoce el epítipo 9E10, EQKLISEEDL, de la proteína humana c-myc (Evan et al., 1985), introducido en el extremo carboxilo de algunas isoformas de 4.1R clonadas.

**Gamma tubulina.** Anticuerpo policlonal anti- $\gamma$ -tubulina (Sigma), derivado de conejo.

**Faloidina-TXR.** Toxina que se une a los polímeros de actina, estabilizándolos. Lleva acoplado isotiocianato de Texas Red (Sigma).

**DM1A.** Anticuerpo monoclonal anti- $\alpha$ -tubulina (Sigma), derivado de ratón.

**AIM-1.** Anticuerpo policlonal (Cell Signaling technology) que reconoce la zona amino terminal de la proteína aurora B.

**Anti-lámina.** Anticuerpo policlonal (Sigma) que reconoce el extremo carboxilo terminal de la lámina B cedido por el Dr. S. D. Georgatos (University of Crete, Greece).

**Anti-anexina II.** Anticuerpo policlonal (Abcam) que reconoce los aminoácidos 9 al 30 de la zona amino de la proteína humana anexina II.

**FLAG.** Anticuerpo policlonal (Sigma) que reconoce el epítipo FLAG.

## 2.2. Anticuerpos secundarios

- Anti-Ig de ratón acoplado al fluoróforo Alexa 488 o 555 (Invitrogen).
- Anti-Ig de conejo acoplado al fluoróforo Alexa 488 o 555 (Invitrogen).
- Anti-Ig de ratón conjugado con peroxidasa de rábano (Amersham Pharmacia Biotech).
- Anti-Ig de conejo conjugado con peroxidasa de rábano (Amersham Pharmacia Biotech).

## 3. Vectores de expresión.

**pSR $\alpha$**  (Takebe et al., 1988). Plásmido de expresión en células de mamífero, con promotor temprano del virus SV40.

**pCR3 y pCR3.1** (Invitrogen). Plásmidos de clonación y expresión, con promotores T7 y Sp6 para sistemas procariotas y promotor temprano de Citomegalovirus (CMV) para células de mamífero.

**pcDNA3.1/CT-GFP-TOPO®** (Invitrogen). Plásmido de clonación y expresión de proteínas de fusión con GFP (proteína verde fluorescente), con el promotor T7 para sistemas procariotas y el promotor temprano de Citomegalovirus (CMV) para expresión en células de mamífero.

#### 4. Oligonucleótidos.

Los oligonucleótidos utilizados en el desarrollo de esta tesis fueron sintetizados por Isogen, Boehringer Mannheim o Boehringer Ingelheim. En la siguiente tabla se muestra la secuencia de cada uno de ellos:

Nombre	Secuencia 5'-3'
FLAG-ATG1- <i>Bgl</i> II	CCCAGCAgATCTATggATTATAAAgATgATgATg ATAAAggATgCACAACAgAgAAgAgTTTAgTgAC
Stop E21- <i>c-myc</i>	CCCgggAgATCTTCATAAgT
Sin E2'	gCCCgggAgATCTCCgCACCCAgCCCgAgAgAAgAgTT
E17-rev- <i>Xho</i> I	CCATTTTCTCgAgTggTCATgTggACAAAgTAC
E17RT-rev	ggTCATGTGGACAAAGTACT
E1A	ggCAggAACCTCTTAAAggg
E1B	gCTTCCATgTgCCACTgCgg
E1C	ggTgCgTTTAgTCAGTCAGC
E2'sense	CAACATCATgACAACAgAgA
E17Stop- <i>Eco</i> RI	TggTACgAATTCTCAGAAgTTCTTTTTCAGCT
<i>Nhe</i> I-red-sense	CCggTCgCTAgCATggTgCgCTCC
Red-Stop- <i>Bgl</i> III- anti	CTAgAgTAgATCTCgCTACAgGAA
Stop E21 <i>Xho</i> I	TTAgTTCTCgAgCTCACTCTACAgCAATCTCgg
E2 <i>Sac</i> I sense	gACTTgACCGAgCTCAAggAgCggACA
E2 225 <i>Sac</i> I	CCCAATTgCAGAgCTCgAACCGgAAC
E2 158 <i>Sac</i> I	gCAGAAACAgAgCTCgCTCaggAAgAAC
Hairpin <i>Xho</i> I sense	TCgAgCCggATCCggCCggATCCggCCggATCCggC
Hairpin <i>Nhe</i> I	CTAgCCCggATCCggCCggATCCggCCggATCCggg
E2A	gCTTTCTgATgTCCgCTCC
E2SBIS	ggAgCggACATCAGAAAgC
E4A	gAAACCTTgCAGTgCATgT
E7A	CTgTTAACTgTgCTgggTCAG
E10S	TggCCTTCTggTTTACAAAg
E11A	CgTgTgATgTTCTACACAg
E16A	ACATTAAATTgCTATgTCTg
E17S	gATTTAgACAAGAgTCAAgAggA
E20A	TTCAATACgTgTCTCTgAA
ATGpreE4 sense	AgCAGTCTCgAgACAATgCCTgCTCaggAAgAACTC
ATGE4mut sense	CCCAAAAACCAATCAGAAAACACaggAACgTgCACTgCAAgg
ATGE4mut anti	CCTTgCAGTgCACgTTCCTgTgTTTTCTgATTggTTTTggg
<i>Bgl</i> III ATGgfp	gCTCgAgAgATCTATggCTAgCAAaggAgAAg
gfp-preE4 <i>Eco</i> RI reverse	AACAggggAATTCTCCAgTAgTgCAAATAAATTTAagg
<i>Eco</i> RI ATG4	CAGAAAACACgAATTCATgCACTgCAAggTTTC

**Tabla 2.** Oligonucleótidos utilizados en el desarrollo de este trabajo.

## Enzimas

### 4.1. Enzimas de restricción.

Se han utilizado durante la clonación de los ADNc de 4.1R y en la subclonación de éstas las siguientes enzimas de restricción: *Bgl*III, *Bam*HI, *Xho*I, *Hind*III, *Eco*RI, *Eco*RV, *Sca*I, *Kpn*I y *Xba*I de las casas Pharmacia Biotech, Boehringer Mannheim, Fermentas, New England Biolabs y Promega.

### 4.2. Enzimas de modificación del ADN, Polimerasas y Retrotranscriptasa.

Se utilizó DNA ligasa del fago T4 de Fermentas. Dentro de las polimerasas termoestables utilizadas se encuentran Taq ADN polimerasa de *Thermus aquaticus* (Promega), *Pfu* Turbo™, de *Pyrococcus furiosus* (Stratagene) y Expand™ High Fidelity PCR System, una combinación de las ADN polimerasas de *Thermus aquaticus* y *Pyrococcus woesei* (Boehringer Mannheim). La retrotranscriptasa utilizada es la del virus de la mieloblastosis aviar (AMV), de Promega.

## 5. Generación de proteínas truncadas, mutadas y de fusión de 4.1R.

En el desarrollo de esta tesis se han utilizado los ADNc para la proteína 4.1R aislados previamente en el laboratorio a partir de células MOLT-4 y subclonados en vectores apropiados que permitiesen bien su expresión en células de mamífero, bien su transcripción y traducción *in vitro*. Estos ADNc han sido utilizados como molde para generar distintas construcciones parciales o quimeras de 4.1R así como mutantes puntuales, de delección o proteínas de fusión a GFP. A continuación se listan todas las construcciones utilizadas y la forma en que se obtuvieron.

### 5.1. Construcción de ADNc parciales y quimeras de 4.1R:

**4.1R80 Δ16, 18, 19.** El ADNc que codifica para esta isoforma de 4.1R se preparó como quimera a partir de las isoformas pSRα 4.1R<sup>80</sup> Δ16 (isoforma traducida a partir del ATG2, con todos los exones excepto el 16) y pSRα 4.1R<sup>135</sup> Δ16, 18, 19. Para ello se cortó el fragmento que va desde la diana *Sca*I hasta la diana *Bam*HI (exón 12 de 4.1) en pSRα 4.1R<sup>80</sup> Δ 16 y se insertó en las dianas equivalentes de pSRα 4.1R<sup>135</sup> Δ 16, 18, 19.

**pCR3 E4-4.1R.** Esta construcción expresa una isoforma de 4.1R con los exones: 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 18, 19, 20, y 21. Se obtuvo mediante la técnica de Reacción en Cadena de Polimerasa (PCR), usando pSRα 4.1R<sup>80</sup> como molde y los oligonucleótidos ATGpreE4 sense y STOPE21-*cm*yc. Se clonó este fragmento en el vector pCR3 en las dianas *Bam*HI y *Bgl*II.

**pCR3 pre-gfpH.** Esta construcción contiene la secuencia posterior al ATG2 del exón 4 y los exones 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19, 20 y 21 que codifican para 4.1R. En la zona anterior al ATG2 se han introducido 157 nucleótidos de la secuencia que codifica para la proteína verde fluorescente GFP. Esta construcción se obtuvo mediante la técnica de la Reacción en Cadena de la Polimerasa (PCR), usando pcDNA3.1/CT-GFP como molde y los oligonucleótidos *Bgl*II ATGgfp y gfp-preE4 *Eco*RI reverse. La digestión del producto de esta reacción con las enzimas *Bgl*II y *Eco*RI fue ligado el vector pCR3 previamente digerido con *Bam*HI y *Eco*RI. Esta construcción se cortó con *Eco*RI y *Xho*I y se ligó con el producto obtenido por amplificación desde la construcción PCR3 E4-4.1R con los oligonucleótidos *Eco*RI ATG4 y Stopc-myc*Xho*I.

## 5.2. Mutagénesis del sitio de iniciación de la traducción en el exón 4 (ATG2).

**pCR3 4.1R<sup>135</sup> ATG2 mut.** Esta construcción se preparó según el kit de mutagénesis dirigida de Stratagene siguiendo las instrucciones del fabricante. Las reacciones se llevaron a cabo utilizando la isoforma pCR3 4.1R<sup>135</sup> como molde y los oligonucleótidos complementarios ATGE4mut sense y ATGE4mut anti, como cebadores.

**pCR3 4.1R<sup>135</sup> Δ16, 19 ATG2 mut.** Esta construcción se preparó de forma análoga a la anterior utilizando como molde pCR3 4.1R<sup>135</sup> Δ 16, 19.

## 5.3. Mutagénesis que introduce un codón de parada de la traducción en el exón 2.

**pCR3 4.1R<sup>135</sup> Δ16, 19 STOP.** Esta construcción se preparó según el kit de mutagénesis dirigida de Stratagene siguiendo las instrucciones del fabricante. Las reacciones se llevaron a cabo utilizando la isoforma pCR3 4.1R<sup>135</sup> como molde y los oligonucleótidos complementarios stop-sense y stop-anti como cebadores.

## 5.4. Construcción bicistrónica DsRed-4.1-EGFP.

Esta construcción se preparó amplificando la secuencia codificante para la proteína DsRed del vector pDsRed1-N1 (Clontech) mediante los oligonucleótidos *Nhe*I-red-sense y Red-Stop-*Bgl*II-anti, que introducen las dianas de restricción *Nhe*I y *Bgl*II en el fragmento amplificado. La digestión de este fragmento con estas enzimas de restricción se purificó y se clonó en el vector pEGFP-N1 (Clontech) previamente digerido con *Bgl*II. Se seleccionaron aquellos clones que en los que DsRed estaba en la orientación correcta. Las secuencias de 4.1R que se quería clonar en este vector se amplificaron mediante la Reacción en Cadena de la Polimerasa (PCR) utilizando como molde el ADNc para 4.1R<sup>135</sup> y los oligonucleótidos específicos. Los productos de estas reacciones se digirieron con la enzima de restricción *Sac*I y se ligaron con el vector bicistrónico digerido con la



misma enzima. Se seleccionaron construcciones con el fragmento de ADN de 4.1 en orientación 5'-3' y 3'-5' mediante digestiones específicas. Los clones positivos fueron secuenciados. Las construcciones DsRed-H4.1s-EGFP y H-DsRed-4.1s-EGFP se realizaron mediante la introducción de una secuencia con alta estructura secundaria que impide el paso del ribosoma (Koromilas et al., 1992) en la zona 5' anterior a la secuencia de 4.1R o en la zona 5' anterior a la secuencia de DsRed mediante la digestión con las enzimas *XhoI* y *NheI*, respectivamente. Se secuenciaron los clones positivos para confirmar la introducción de la secuencia con alta estructura secundaria.

## 6. Secuenciación de los ADNc.

Se llevó a cabo mediante el método de terminadores marcados con fluorocromos y PCR, usando los reactivos suministrados en el ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit, con AmpliTaq® DNA Polimerasa, FS (Perkin-Elmer), siguiendo siempre las recomendaciones del fabricante. Las reacciones de PCR se llevaron a cabo en un termociclador GeneAmp 9600 (Perkin-Elmer) y fueron analizadas en los secuenciadores automáticos ABI 373A o ABI PRISM 377 (Perkin-Elmer). Las secuencias de los oligonucleótidos utilizados para las reacciones se detallan en la tabla 1.

## 7. Transcripción y traducción *in vitro*.

Para los experimentos de transcripción y traducción *in vitro* a partir de ADNc, se utilizó el sistema TNT® Couple Reticulocyte Lysate System (Promega). El molde es en todos los experimentos un ADNc clonado en el vector pCR3.1 y el marcaje de la proteína sintetizada se realizó con PROMIX™; una mezcla de <sup>35</sup>S L-Metionina y <sup>35</sup>S L-Cisteína, con una actividad de 14,3 mCi/ ml (Amersham Pharmacia Biotech). Las reacciones se llevaron a cabo según las instrucciones del fabricante.

## 8. Transfección transitoria de ADNc en células de mamífero.

La transfección transitoria de ADNc en células COS-7, HeLa y MDCK se realizó mediante electroporación con un electroporador Electro Cell Manipulator 600 (BTX). Las células fueron tripsinizadas, centrifugadas y resuspendidas en un volumen final de 200 µl de medio DMEM suplementado con suero fetal de ternera 10% (v/v), glutamina 1% (p/v), penicilina 50 u/ml y estreptomicina 50 u/ml, 10mM HEPES (Sigma) y 37,5 mM NaCl. Posteriormente se mezclaban con 1-5 µg del ADN plasmídico a transfectar y con 20 µg de espermatozoides de salmón (Sigma-Aldrich). Tras un pulso eléctrico de 200 V, 480 Ω y 950 µF, las células se volvían a sembrar en placas y se procesaban a las 24 o 48 horas posteriores a la transfección.

## 9. Obtención de extractos proteicos.

Para obtener extractos de proteínas totales, las células fueron recogidas, lavadas dos veces con PBS (solución salina tamponada con fosfato, 8 mM Na<sub>2</sub>HPO<sub>4</sub>, 1,5 mM KH<sub>2</sub>PO<sub>4</sub> pH 7,2; 137 mM NaCl y 2,7 mM KCl) y lisadas con solución de Laemmli (Laemmli, 1970). Posteriormente fueron sonicadas durante 1 minuto para romper el ADN celular.

## 10. Aislamiento de centrosomas de células MOLT-4.

Los centrosomas humanos se aislaron de células MOLT-4 siguiendo el protocolo descrito por Moudjou y Bornens en 1998. Se eligieron estas células porque dicho protocolo está optimizado para células T humanas. Brevemente, las células se trataron con nocodazol y citocalasina D antes del aislamiento de sus centrosomas. La suspensión de centrosomas obtenida se depositó en gradientes discontinuos de sacarosa siguiendo un protocolo descrito previamente (Moudjou y Borners, 1998). Para el análisis mediante inmunofluorescencia las fracciones del gradiente se depositaron en un cubreobjetos y se testaron con los anticuerpos anti- $\gamma$ -tubulina o con anti-4.1R mediante inmunofluorescencia.

## 11. Electroforesis de proteínas (SDS-PAGE) y electrotransferencia (*Western Blot*).

Las proteínas se separaron y analizaron mediante electroforesis en geles de poliacrilamida (SDS-PAGE) preparados según el sistema de geles discontinuos (Laemmli, 1970). Las muestras se prepararon en solución Laemmli y se hirvieron durante 5 minutos antes de analizarlas en condiciones reductoras en minigeles adaptados al sistema de electroforesis Mini-PROTEAN II (Bio-Rad). Como marcadores se utilizaron los SDS-6H, una mezcla de 6 proteínas de pesos moleculares comprendidos entre 205 y 29 kDa, de Sigma Aldrich. Los geles se corrieron a un voltaje constante de 200V durante 1 hora, en una solución tamponada con 25 mM Tris, 130 mM glicina y SDS 0,1% (p/v).

En el caso de marcaje radiactivo de las proteínas, los geles se fijaron en una solución de etanol al 20% (v/v) y ácido acético al 10% (v/v) durante 20 minutos. A continuación se tiñeron con una solución compuesta por 2,5 g/l de Coomassie Brilliant Blue (Boehringer Ingelheim), metanol al 40% (v/v) y ácido acético al 10% (v/v) durante 30 minutos. Se destiñeron eliminando el colorante no unido con una solución de metanol al 45% (v/v) y ácido acético al 7%. Los geles se secaron y se expusieron para autorradiografía (Kodak).

Para inmunoelectrotransferencia, las proteínas separadas por SDS-PAGE se transfirieron a una membrana de Immobilon®-PVDF (Millipore) en una solución de Tris-Borato pH 8,2 (0,25 M ácido bórico y 0,25M Tris/HCl), usando el sistema Mini Trans-Blot (Bio-Rad). Las transferencias se realizan durante 1 hora a un voltaje de 100V. Posteriormente la membrana se procesó como se describe en protocolos publicados con anterioridad (De Cárcer et al., 1995). La

detección de los complejos antígeno-anticuerpo sobre las membranas se realizó mediante una reacción quimio-luminiscente catalizada por la peroxidasa de rábano unida a los segundos anticuerpos, usando los reactivos ECL (Enhanced ChemiLuminiscence) de Amersham Pharmacia Biotech y película fotográfica (Kodak).

## 12. Microscopía.

Para inmunomicroscopía de fluorescencia las células adherentes crecieron sobre cubreobjetos y las células MOLT-4 se incubaron sobre cubreobjetos recubiertos con 1mg/ml de poli-Lisina (Sigma-Aldrich) en placas de pocillos de fondo plano (Costar Corp., Cambridge, MA). Se han utilizado dos tipos de fijaciones, bien con formalina bien con metanol. Para la primera fijación las células se incubaron con formalina (Formaldehído al 37%, Sigma Aldrich) al 10% (v/v) en PBS durante 20 minutos a temperatura ambiente. El exceso de formaldehído se bloqueó incubando 10 minutos con 10 mM glicina. Las células así fijadas se permeabilizaban posteriormente con Tritón X-100, al 0,2% (v/v) en PBS. La fijación mediante metanol se llevaba a cabo incubando las células a -20°C en metanol puro (MERCK) durante 5 minutos.

Después de esto, se incubaron las células con los anticuerpos apropiados según los protocolos ya publicados (De Cárcer et al., 1995). El ADN celular se teñía incubando las células con DAPI (4', 6'- diamidino-2-fenilindol, de Sigma-Aldrich), 1 mg/ml durante 5 minutos a temperatura ambiente.

Las preparaciones así tratadas fueron analizadas con un microscopio de epifluorescencia Zeiss Axioskop (Carl Zeiss) o con un sistema confocal Radiance 2000 de Bio-Rad acoplado a un microscopio invertido de Zeiss modelo Axiovert S100 TV. Las imágenes obtenidas fueron procesadas con el programa Confocal-Assistant o Adobe-Photoshop.

## 13. Generación de anticuerpos policlonales.

El péptido KTGEGVSTLST, en el que se incluye la zona carboxilo terminal característica de las isoformas de 4.1R sin CTD, fue obtenido en un sintetizador automático de multipéptidos (AMS 422, Abimed, Langerfeld, Germany) utilizando un procedimiento de fase sólida y química F-moc estándar. Los péptidos fueron acoplados a KLH (Keyhole Limpet Hemocyanin) y se inmunizó con ello a dos conejos. Se recogió tanto el suero preinmune de ambos conejos, como los antisueros.

## 14. Inmunoprecipitaciones.

Las células se lavaron dos veces con PBS y se levantaron en PBS con EDTA (5mM). Se incubaron en tampón de lisis (10mM Tris/HCl (pH 7.6), 150 mM NaCl, 1% Tritón Tx-100, 5 mM EDTA, 0,5 mM PMSF y 1 µg/ml de leupeptina, aprotinina y peptasina). Se centrifugó en una minifuga a 4°C durante 10 minutos. El sobrenadante se incubó con el anticuerpo utilizado para

inmunoprecipitar asociado a sefarosa durante 3 h a 4°C y se procesó siguiendo protocolos descritos previamente (Lallena and Correas, 1997).

### **15. Centrifugación por velocidad de sedimentación en gradientes de densidad en sacarosa.**

Las células se lavaron, lisaron y sedimentaron siguiendo protocolos ya descritos (Echeverri et al., 1996). Brevemente, las células se lisaron en un tampón que contenía 50mM Tris/HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% Nonidet p40, 0,5 mM PMSF y 1 µg/ml de leupeptina, aproptinina y pepstatina, manteniéndose 15 minutos a 4°C. Los lisados se centrifugaron a 100 000 g durante 25 minutos a 4°C en una ultracentrifuga Beckman TL-100 usando el rotor TLA-100. Los sobrenadantes se cargaron sobre 4,2 ml de un gradiente del 5-20% (peso/volumen) de sacarosa. Los gradientes con las muestras se centrifugaron en un rotor TST-60.4 (Kontron Instruments) a 26 000 rpm durante 18 h a 4°C y se recogieron fracciones de 300 µl. Los controles de sedimentación, procesados en paralelo fueron BSA (4,4S), catalasa (11,3 S) y tiroglobulina (19S). Volúmenes idénticos de cada fracción se analizaron mediante SDS-PAGE y *Western Blot*.

### **16. Aislamiento de ARN total de células en cultivo.**

El aislamiento del ARN se realizó en células COS-7 transfectadas con diferentes construcciones. Se utilizó Trizol (GibcoBrl). Para ello se siguieron las instrucciones detalladas en el protocolo de dicho reactivo. Se añadió 1 ml de Trizol a una placa P-100 de células. Se incubaron durante cinco minutos a temperatura ambiente. Se añadió 0,2 ml de cloroformo/ml de trizol. Se agitó bien la mezcla y se centrifugó a 12.000 x g durante 15 minutos y a 4°C. Se recogió la fase acuosa y se precipitó con isopropanol (0,5ml/ ml Trizol). Se incubaron las muestras a temperatura ambiente durante 10 minutos, posteriormente se centrifugó a 12.000 x g durante 10 minutos a 4°C. Se recogió el sobrenadante y se lavó una vez en etanol al 75% (v/v). Finalmente se centrifugó y se eliminó el sobrenadante. El sedimento se dejó secar 10 minutos, y se resuspendió en agua-DEPC (dietil pirocarbonato).

### **17. Transferencia del ARN a membrana e hibridación (*Northern Blot*).**

Se sometió a electroforesis a 60V en el tampón MOPS (ácido 3-(N-morpholino) 20 mM propanosulfónico pH 7, Acetato sódico 8 mM y EDTA 1 mM pH 8) geles de agarosa al 0.8% con formaldehído al 6% (p/v) en los que se había cargado 5µg del ARN sometido a análisis. Se llevó a cabo la transferencia a membranas de nylon (Amersham Biosciences) según la técnica de transferencia capilar descrita previamente (Alwine et al., 1977). Se marcó el ADN utilizado como sonda con ( $\alpha$ -<sup>32</sup>P) dCTP mediante el sistema de marcaje de sondas rediprime™ II (Amersham farmacia biotech), y siguiendo las instrucciones descritas por el fabricante. La hibridación se llevó a cabo con el tampón de hibridación ExpressHyb (Clontech) a 68°C durante 12 horas. Las membranas se

lavarón dos veces con 2 x SSC (1 x SSC: 150 mM NaCl, 15mM citrato sódico, pH 7.0 ) 0,1% SDS a 42°C durante 20 minutos y dos veces en 0,1 x SSC 0,1% SDS a 56 °C durante 30 minutos. Se expuso la membrana a una película de autorradiografía a -80°C durante 12h.

## **18. Clonación de los ADNc de la proteína 4.1R sin dominio CTD.**

### **18.1. Retrotranscripción.**

Como material de partida se usó un ARN total comercial (BD Biosciences) obtenido de cerebro humano. Sobre este molde se realizó una reacción de retrotranscripción (RT), usando como cebador el oligonucleótido E17RTrev, y la retrotranscriptasa de AMV. La reacción se llevó a cabo en un volumen final de 20 µl, 2h a 42°C, en un termociclador GeneAmp PCR System 9600 (Perkin-Elmer). Previamente se desnaturalizó el ARN molde, incubándolo durante 2 minutos a 85°C y un minuto a temperatura ambiente.

### **18.2. Reacción en Cadena de la Polimerasa (PCR).**

El producto de la retrotranscripción consiste en una mezcla de ADNc específicos de la proteína 4.1R sin carboxilo terminal. Sobre el mismo, se realizaron dos amplificaciones por PCR, usando como cebadores los oligonucleótidos SinE2' y E17RTrev para una de ellas y E2'sense y E17RTrev para la otra. El volumen de reacción final fue de 25 µl y las condiciones: 4 minutos a 94°C, seguidos de 35 ciclos de 30 segundos a 94°C, 30 segundos a 50°C y 3 minutos a 68°C; para completar la elongación de las cadenas de ADNc y favorecer la introducción de adeninas extras en el extremo 3' se mantuvo la reacción 5 minutos a 72°C. La reacción se llevó a cabo en un termociclador GeneAmp PCR System 9600 (Perkin Elmer).

### **18.3. Clonación de los ADNc amplificados.**

Los ADNc productos de la amplificación fueron purificados usando geles de agarosa y GeneClean®II (BIO 101), para posteriormente ser clonados en el vector pcDNA3.1/CT-GFP-TOPO® (Invitrogen) mediante apareamiento por las timinas de este vector con las adeninas del extremo 3' de los ADNc siguiendo las instrucciones del fabricante.

### **18.4. Subclonación de los ADNc.**

Los ADNc clonados directamente en el vector pcDNA3.1/CT-GFP-TOPO fueron subclonados en el vector de expresión eucariota pCR3.1. Para ello se amplificó el ADNc con los cebadores Flag-ATG1-BglII y E17RT-rev-XhoI. Esto permitió introducir en el extremo 5' de estos ADNc el epítipo FLAG. Este producto de PCR se purificó y se cortó con las enzimas XhoI y BglII para poder ligarlo en el vector pCR3.1.

### **18.5. Experimentos de citometría de flujo.**

Las células se tripsinizaron y lavaron con PBS. Se resuspendieron en 1 ml de PBS y se analizaron en un equipo de citometría de flujo analítica FACSCalibur (BD Biosciences) con el programa Cell Quest-Pro.

## Resultados



## **RESULTADOS**

Los mecanismos de procesamiento alternativo del pre-ARNm de 4.1R parecen ser los máximos responsables de la generación de la gran diversidad de isoformas de la proteína 4.1R. En este trabajo se describe otro mecanismo responsable de la generación de esta diversidad y se aíslan nuevas isoformas con un patrón de procesamiento característico.

### **1. GENERACIÓN DE DOS ISOFORMAS DE 4.1R DESDE LOS ADNc QUE CONTIENEN EL EXÓN 2'.**

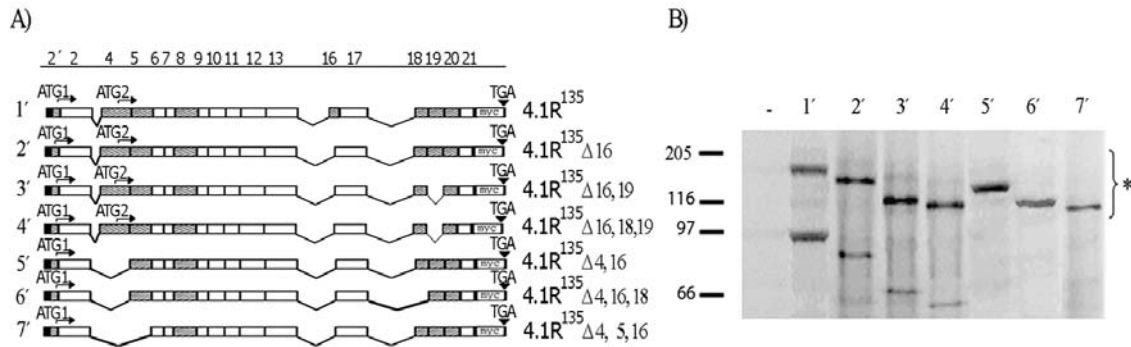
#### **1.1. Algunos ADNc para 4.1R<sup>135</sup> originan dos proteínas *in vitro*.**

Con el fin de estudiar la expresión *in vitro* de diferentes ADN codificantes (ADNc) para isoformas de 4.1R se llevaron a cabo reacciones de transcripción y traducción *in vitro* a partir de ellos. Estas reacciones se realizaron en lisados de reticulocitos de conejo utilizando el sistema TNT® Coupled Reticulocyte Lysate System. Este sistema utiliza ARN polimerasa del fago T7 para la transcripción. Los moldes para estas reacciones fueron distintos ADNc para isoformas concretas de 4.1R aislados previamente en el laboratorio a partir de ARN citoplasmático de la línea linfóide humana MOLT-4 y clonados en los vectores de expresión pCR3.1. Un esquema de cada uno de estos ADNc se muestra en la figura 16A. A la derecha de cada construcción se muestra su nombre según la composición exónica que éstos tienen. Dado que dicha nomenclatura es compleja, se ha denominado a cada uno de los ADNc con un número (1', 2', 3', 4', 5', 6' y 7'), con el que nos referiremos a ellos de forma abreviada a lo largo de este trabajo. Estos ADNc contienen el exón 2' y por tanto el ATG1 que dará lugar a la síntesis de isoformas de 4.1R de peso molecular alto (4.1R<sup>135</sup>).

Los resultados de las reacciones de transcripción y traducción *in vitro* de los distintos ADNc se muestran en la figura 16B. Como se observa en la autorradiografía, todos los ADNc dan lugar a una proteína del tamaño esperado (150, 140, 135, 125, 135, 115 y 100 kDa, respectivamente) señaladas con una llave en la figura. Inesperadamente, en la transcripción-traducción *in vitro* de cuatro de estos ADNc (1', 2', 3' y 4') aparece además una segunda banda de menor tamaño.

Comparando la composición exónica de los siete ADNc utilizados se observa que dos de ellos, el 5' (4.1R<sup>135</sup> Δ4, 16) y el 2' (4.1R<sup>135</sup> Δ16) difieren únicamente en la ausencia o presencia del exón 4, respectivamente. El primero de estos ADNc, carece del exón 4, y da lugar a una proteína del tamaño esperado. El segundo, contiene el exón 4 y genera en la reacción *in vitro* dos proteínas. El resto de ADNc que genera dos proteínas en la transcripción y traducción *in vitro* contienen también el exón 4 (1', 2', 3' y 4') mientras que todos aquellos ADNc que carecen de dicho exón (5', 6' y 7') producen únicamente una proteína. Esto sugiere que el exón 4 es necesario para la producción *in vitro* de la proteína de menor tamaño.





**Figura 16. Proteínas sintetizadas *in vitro* a partir de ADNc para isoformas de 4.1R<sup>135</sup>.** A) Esquema de la composición exónica de los diferentes ADNc para isoformas 4.1R<sup>135</sup>. Los exones se han representado como sigue: alternativos, en gris; constitutivos, en blanco; no codificantes, en negro. El número de cada exón se indica en la parte superior de los esquemas. Se indican los dos codones de iniciación de la traducción, situados en los exones 2' y 4 (ATG1 y ATG2 respectivamente); así como el codón de parada de la traducción (TGA), situado en el exón 21. A la derecha de cada construcción se incluye el nombre que recibe cada una de ellas en función de su composición exónica y a la izquierda un número, con el que las nombraremos a lo largo de este trabajo (1', 2', 3', 4', 5', 6' y 7'). B) Autorradiografía de las reacciones de transcripción y traducción *in vitro*, en presencia de <sup>35</sup>S-Metionina, de los ADNc que se muestran en el apartado A. El primer carril (-) corresponde a una reacción de transcripción y traducción *in vitro* que carecía del molde de ADNc. Los carriles numerados del 1' al 7' se corresponden con las reacciones de transcripción y traducción *in vitro* partir de los ADNc numerados de igual forma. Se han señalado con una llave las proteínas que presentan la movilidad electroforética esperada para proteínas traducidas desde el ATG1 de cada uno de los ADNc.

## 1.2. Las proteínas de menor tamaño sintetizadas *in vitro* desde ADNc para 4.1R<sup>135</sup> presentan el mismo tamaño que las isoformas correspondientes de 4.1R<sup>80</sup>.

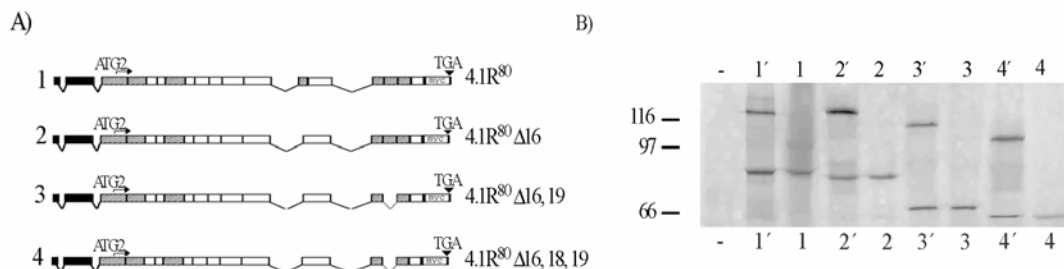
Además de los ADNc para isoformas 4.1R<sup>135</sup> existen ADNc para otro grupo de isoformas de 4.1R de peso molecular intermedio, unos 80 kDa, a las que denominaremos 4.1R<sup>80</sup>. Estos últimos ADNc carecen del exón 2' en el que se localiza el codón de inicio de la traducción (ATG1). Esto hace que la traducción de estos ADNc se realice desde un codón de inicio de la traducción situado en el exón 4 (ATG2).

En el laboratorio existen distintos ADNc para 4.1R<sup>80</sup>, (aislados de ARN citoplasmático de la línea linfocítica humana MOLT-4, y clonados en el vector de expresión pCR3). Se eligieron aquellos que tenían una composición exónica idéntica a la de los ADNc 1', 2', 3' y 4', salvo por la ausencia del exón 2'. Estos distintos ADNc para 4.1R<sup>80</sup> se nombraron 1, 2, 3 y 4 respectivamente y su composición exónica concreta se muestra en la figura 2A. Se llevó a cabo la transcripción y traducción *in vitro* de estos ADNc y de los ADNc 1', 2', 3' y 4'.

Como puede observarse en la figura 17B, las reacciones *in vitro* de los ADNc 1, 2, 3 y 4 dan lugar a una única proteína con el tamaño esperado para cada una de las isoformas: 82, 80, 72 y 65 kDa, respectivamente. Todos los productos proteicos de menor tamaño resultantes de la reacción *in vitro* de

ADNc para las isoformas 4.1R<sup>135</sup> presentan la misma movilidad electroforética que las proteínas resultantes de la transcripción y traducción *in vitro* de los ADNc para las isoformas de 4.1R<sup>80</sup>.

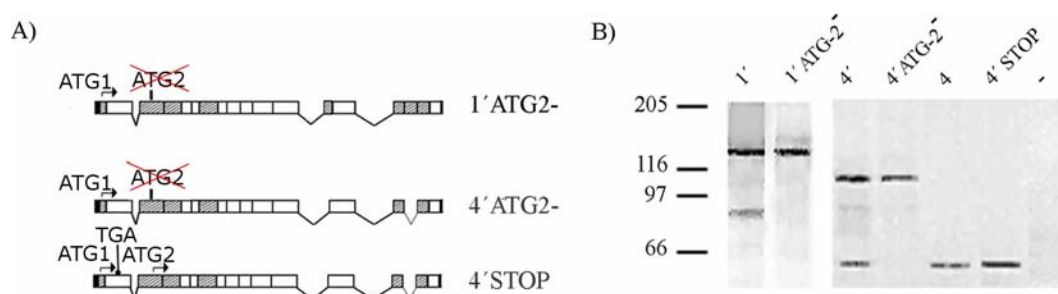
Estos resultados sugieren que las proteínas de menor tamaño obtenidas en la transcripción y traducción *in vitro* de los ADNc para 4.1R<sup>135</sup> podrían ser traducidas a partir del ATG2 presente en el exón 4 y que, por lo tanto, serían también isoformas de 4.1R.



**Figura 17.** Una de las dos proteínas generada *in vitro* a partir de un ADNc para 4.1R<sup>135</sup> presenta la misma movilidad electroforética que la proteína que se obtiene de la transcripción y traducción *in vitro* del correspondiente ADNc para la isoforma 4.1R<sup>80</sup>. **A)** Esquema de la composición exónica de los ADNc para 4.1R<sup>80</sup> utilizados para su transcripción y traducción *in vitro*. Su composición se ha indicado de igual forma que en la figura 16. A la derecha de cada ADNc se muestra el nombre que recibe éste atendiendo a su composición exónica y a la izquierda un número con el que las nombraremos a lo largo de este trabajo (1, 2, 3 y 4). **B)** Autorradiografía de las reacciones de transcripción y traducción *in vitro*, en presencia de <sup>35</sup>S-Metionina, de los ADNc que se muestran en el apartado A y de los ADNc 1', 2', 3' y 4'. El primer carril (-) corresponde a los productos obtenidos en una reacción de transcripción y traducción *in vitro* que carecía del molde de ADNc. Los carriles se han nombrado con el número del ADNc que se ha utilizado para la reacción de transcripción y traducción *in vitro*.

### 1.3. Las proteínas de menor tamaño obtenidas a partir de ADNc de 4.1R<sup>135</sup> son isoformas de 4.1R traducidas desde el ATG2.

Para comprobar si la proteína de menor tamaño sintetizada *in vitro* a partir de ADNc para isoformas de 4.1R<sup>135</sup> se estaba traduciendo realmente desde el ATG2 del exón 4 se generaron construcciones mutantes. Los ADNc 1 y 4 se sometieron a mutagénesis dirigida para sustituir el codón de iniciación del exón 4, ATG, por GTG. Estas construcciones se denominaron 1'ATG2- y 4'ATG2- (fig. 18A). La transcripción y traducción *in vitro* de estas dos construcciones llevadas a cabo a la vez que sus correspondientes ADNc silvestres 1' y 4' (fig. 16A) dieron lugar a una sola proteína correspondiente a una isoforma de peso molecular alto (fig. 18B). Estos resultados indican que las proteínas de menor tamaño obtenidas en la transcripción y traducción *in vitro* de los ADNc 1', 2', 3' y 4' son isoformas de 4.1R traducidas desde el ATG2.



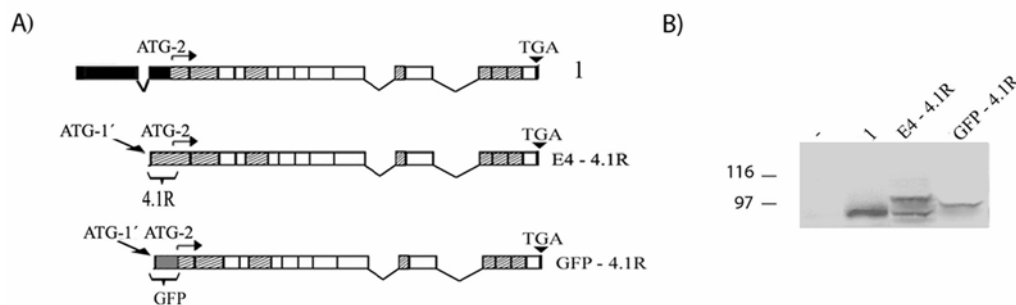
#### 1.4. Las isoformas de 4.1R<sup>80</sup> obtenidas a partir ADNc para 4.1R<sup>135</sup> no son consecuencia de una proteólisis.

Los resultados anteriores sugerirían que las isoformas de 4.1R<sup>80</sup> obtenidas a partir de los ADNc para 4.1R<sup>135</sup> no son consecuencia de una proteólisis de las isoformas de 135 kDa. No obstante, cabe la posibilidad de que la mutación del ATG2 modifique un sitio de proteólisis. Para descartar esta posibilidad se sustituyó mediante mutagénesis dirigida el nucleótido 191 del exón 2, adenina, por timina, en el ADNc 4'. Este cambio produce un codón de parada de la traducción (TAG) en lo que sería el aminoácido 68 de la proteína silvestre. Este ADNc mutante se denominó 4' STOP y un esquema de su composición exónica se muestra en la figura 18A. La traducción *in vitro* de esta construcción no originó la proteína de peso molecular alto, como cabía esperar (comparar carril 4' y 4' STOP fig. 18B). Sin embargo, se sigue generando una proteína del tamaño esperado para la isoforma de peso molecular intermedio (comparar carril 4 y 4' STOP fig 18B). Por lo tanto no es necesaria la producción de la isoforma de 4.1R de peso molecular alto para que se genere la de peso molecular intermedio. Esto indica que la isoforma de 4.1R de peso molecular intermedio no se produce por proteólisis de la isoforma de peso molecular alto.

#### 1.5. La región 5' anterior al ATG2 de 4.1R es suficiente para dirigir la síntesis de la proteína 4.1R<sup>80</sup>.

Todos los ADNc para 4.1R<sup>135</sup> que dan lugar a la isoforma de peso molecular intermedio contienen no sólo el exón 4 sino también el exón 2' y el 2. Se planteó si la presencia del exón 2' y 2 era necesaria para la generación de la isoforma de peso molecular intermedio.

Para comprobar esto se hicieron dos construcciones quimeras cuyos esquemas se muestran en la figura 19. Una de ellas, E4-4.1R, es un ADNc similar al ADNc 1' pero que carece del exón 2' y 2. La secuencia de 4.1R que porta E4-4.1R comienza en el exón 4. Además de esto, el primer triplete del exón 4 se mutó para introducir un codón de iniciación de la traducción ATG1' delante del ATG2. A este codón de iniciación lo denominaremos ATG1'. La transcripción y traducción *in vitro* de esta construcción dio lugar a dos proteínas (fig. 19A). La de mayor peso molecular presenta el tamaño esperado para la proteína traducida desde el ATG1', mientras que la de menor tamaño presenta la misma movilidad electroforética que la que se obtiene en la reacción *in vitro* del ADNc 1 (comparar 1 y E4-4.1R en fig. 19B). Esto indica que, *in vitro*, los exones 2 y 2' no son imprescindibles para la utilización del ATG2, ya que una construcción carente de estos exones es capaz de producir dos proteínas.

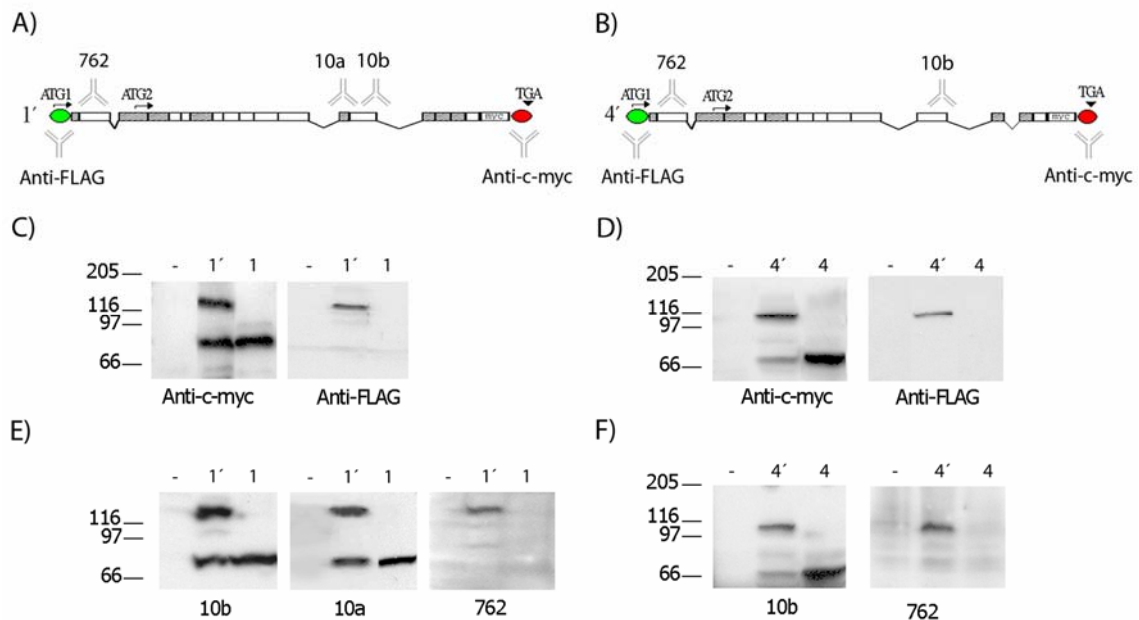


**Figura 19. La secuencia 5' anterior al ATG2 es necesaria para la síntesis de la isoforma de 4.1R de 80 kDa.** A) Esquema de los ADNc para 4.1R<sup>80</sup> usados en la transcripción y traducción *in vitro* que se muestran en la autorradiografía del apartado B. La primera construcción se corresponde con un ADNc para una isoforma natural de 4.1R<sup>80</sup> (1), la segunda (E4-4.1R) es una construcción mutante que carece del exón 2 e introduce un codón de iniciación de la traducción (ATG1') en el extremo 5' del exón 4. En la tercera construcción (GFP-4.1R) la secuencia de 4.1R comprendida entre el ATG1' y el ATG2 ha sido sustituida por un fragmento de la secuencia codificante para GFP. B) Autorradiografía de las reacciones de transcripción y traducción *in vitro*, en presencia de <sup>35</sup>S-Metionina, de los ADNc mostrados en el apartado A. El primer carril de esta autorradiografía (-) muestra un control de la reacción de transcripción y traducción *in vitro* sin ADNc. El ADNc utilizado para cada una de las reacciones de transcripción y traducción *in vitro* se indica en cada uno de los carriles de la autorradiografía.

Para testar la necesidad de la secuencia del exón 4 anterior al ATG2 se realizó otra construcción quimera, a la que se denominó GFP-4.1R. Esta construcción es similar a la construcción E4-4.1R con la diferencia de que la secuencia de 4.1R que va desde el ATG1' hasta el ATG2 del exón 4 fue sustituida por otra secuencia que no codifica para 4.1R (fig. 19A). Esta construcción quimera y el ADNc 1 se transcribieron y tradujeron *in vitro*. La construcción quimera GFP-4.1R produce una única proteína (comparar E4-4.1R y GFP-4.1R fig. 19B). Estos resultados sugieren que la secuencia anterior al ATG2 es necesaria para generar la isoforma de 4.1R de 80 kDa.

### 1.6. Expresión *in vivo* de ADNc para isoformas 4.1R<sup>135</sup>.

Los experimentos anteriores se desarrollaron mediante transcripción y traducción *in vitro* de los ADNc de 4.1R en lisados de reticulocitos. Se planteó a continuación si la obtención de dos isoformas de 4.1R se producía también en experimentos de transfección en células en cultivo. Aunque habría sido interesante realizar un estudio del comportamiento de las isoformas endógenas de 4.1R, la gran variabilidad de las mismas hace imposible este tipo de abordaje. Para solventar este problema se realizaron transfecciones de los cuatro ADNc que habían generado *in vitro* dos isoformas de 4.1R (1', 2', 3' y 4') a los que se fusionaron los epítomos FLAG y c-myc en los extremos 5' y 3', respectivamente. Esto permitió distinguir las isoformas generadas a partir del ADNc transfectado de las isoformas de 4.1R endógenas de la célula.



**Figura 20. Expresión *in vivo* de los ADNc que generan dos isoformas de 4.1R en los estudios *in vitro*.** A y B) Esquema del ADNc 1' y 4' respectivamente a los que se han añadido los epítomos FLAG (en verde) y c-myc (en rojo) en los extremos 5' y 3' respectivamente. Se han señalado las zonas que codifican para los epítomos de los diferentes anticuerpos utilizados (anti-FLAG, anti-c-myc, 762, 10a y 10b). C) Western Blot con el anticuerpo anti-c-myc y anti-FLAG de extractos proteicos totales de células COS-7 no transfectadas (-), transfectadas con el ADNc 1' representado en A o transfectadas con el ADNc 1. D) Western Blot con el anticuerpo anti-c-myc y anti-FLAG de extractos proteicos totales de células COS-7 no transfectadas (-), transfectadas con el ADNc 4' representado en A o transfectadas con el ADNc 4. E) Western Blot con diferentes anticuerpos anti-4.1R (10b, 10a y 762) de células COS-7 no transfectadas; transfectadas con el ADNc 1' representado en A o transfectadas con el ADNc 1. F) Western Blot con diferentes anticuerpos anti-4.1R (10b y 762) de células COS-7 no transfectadas; transfectadas con el ADNc 4' representado en A o transfectadas con el ADNc 4.

Se transfectaron células COS-7 con estas construcciones y se analizaron extractos de las mismas mediante Western Blot con los anticuerpos anti-FLAG y anti-c-myc; así como con anticuerpos anti-4.1R que reconocen la



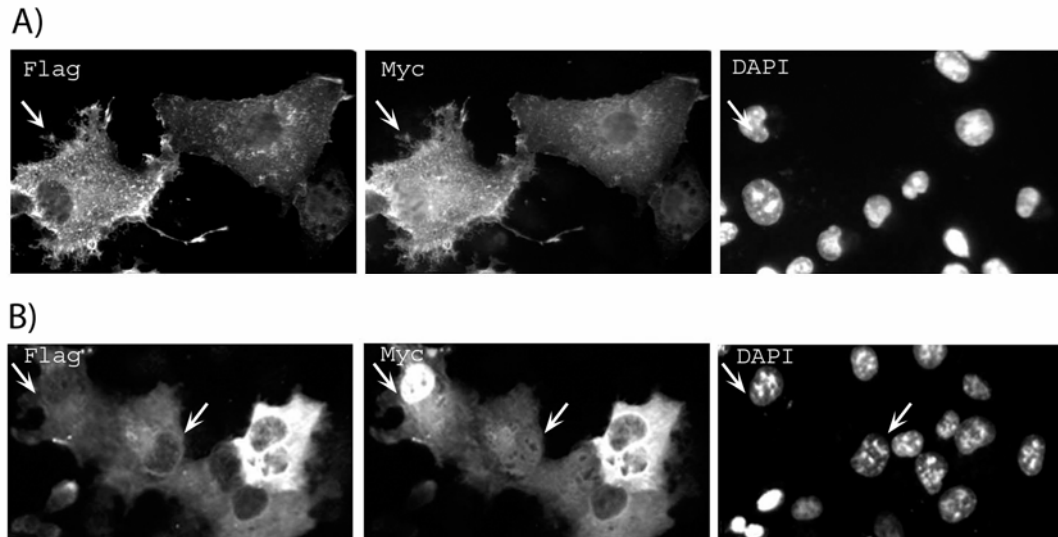
zona amino de la proteína (762) y la zona carboxilo de la misma (10b y 10a). En la figura 20A se muestra un esquema de la zona que reconocen estos anticuerpos. En todos los experimentos se transfectó también el ADNc para las isoformas correspondientes de 80 kDa (ADNc 1, 2, 3 y 4) como control. La figura 20 muestra el resultado de estos *Western blots* para los ADNc 1' (izquierda) y 4' (derecha). El análisis de los extractos de células transfectadas con estos ADNc da lugar a dos proteínas detectadas por los anticuerpos que reconocen la zona carboxilo de la proteína como son anti-c-myc y los anticuerpos anti-4.1 10b y 10a. Esto indica que desde un único ADNc se pueden expresar dos proteínas y que ambas son isoformas de 4.1R.

De estas dos proteínas solo la de mayor peso molecular es detectada por los anticuerpos que reconocen la zona amino, como son los anticuerpos anti-FLAG y el anticuerpo anti-4.1 762 (comparar carriles 1' en fig. 20C y carriles 4' en figura 20D). El tamaño que presenta la proteína que es reconocida tan solo por los anticuerpos anti-carboxilo (anti-myc, 10a y 10b) es el mismo que tienen las isoformas de 4.1R de 80 kDa producidas desde los ADNc 1 y 4 (comparar 1' y 1 en fig. 20C y 20E; y 4' y 4 en fig. 20D y 20F). Estos resultados sugieren que la proteína de menor tamaño expresada *in vivo* es una isoforma de 4.1R que se traduce desde el ATG2.

#### 1.7. Análisis de la expresión *in vivo* de diferentes isoformas de 4.1R<sup>135</sup> mediante inmunofluorescencia.

Con el fin de analizar la expresión *in vivo* se transfectaron también, células COS-7 con los ADNc para 1' y 4' marcados con los epítomos FLAG y c-myc en los extremos 5' y 3' respectivamente para su análisis mediante inmunofluorescencia con los anticuerpos anti-FLAG y anti-c-myc. Como se muestra en la figura 21, hay células que no presentan la misma tinción con los dos anticuerpos (anti-FLAG y anti-c-myc). Existen células en las que la proteína expresada se localiza en el citoplasma cuando es detectada con el anticuerpo anti-FLAG mientras que esa misma célula presenta tinción citoplasmática y nuclear cuando es analizada con el anticuerpo anti-c-myc. Las células con una tinción diferente para los anticuerpos anti-c-myc y anti-FLAG han sido señaladas en la inmunofluorescencia con una flecha. Esta tinción diferencial no ocurre en todas las células, ya que algunas presentan la misma tinción para el anticuerpo anti-FLAG que para el anticuerpo anti-c-myc, lo que impide determinar si se están expresando las dos isoformas de 4.1R. Estos experimentos sugieren que la tinción nuclear se debe a la expresión en algunas células de la isoforma de peso molecular intermedio. Estos datos concuerdan con los descritos previamente en nuestro grupo (Luque et al., 1999).

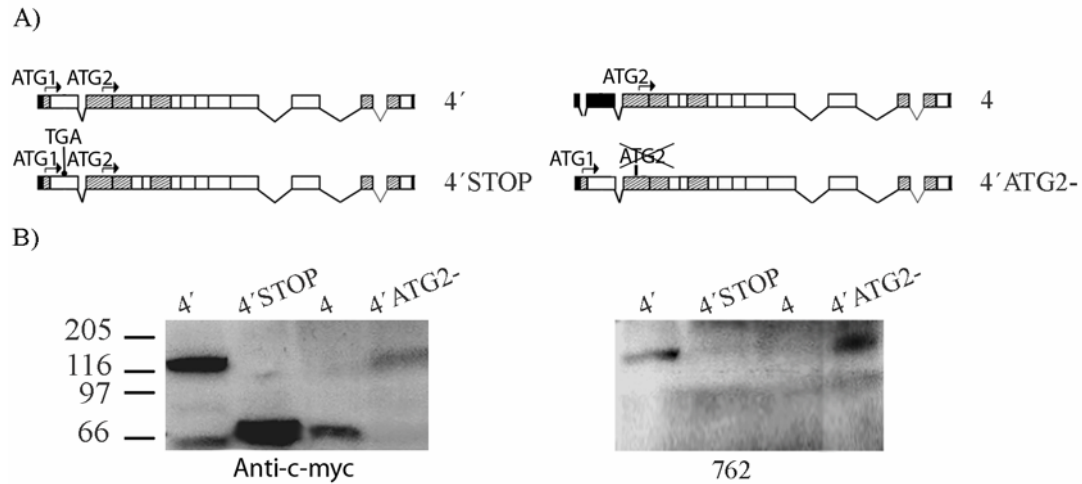
Estos resultados sugieren que la proteína de menor tamaño expresada *in vivo* es una isoforma de 4.1R que se traduce desde el ATG2.



**Figura 21. Análisis de la expresión *in vivo* de los ADNc 1' y 4' en células COS-7 mediante inmunofluorescencia.** **A)** Se transfectaron células COS-7 con el ADNc 1' que porta las secuencias codificantes para los epítopos FLAG y c-myc en los extremos 5' y 3' respectivamente. Se analizaron mediante inmunofluorescencias a las 48 h con los anticuerpos anti- FLAG (panel de la izquierda) y con el anticuerpo anti-c-myc (panel central). La tinción nuclear con DAPI (4'-6'-diamidino-2-fenilindol) se muestra en el panel de la derecha. Se ha señalado con flechas algunas células con tinción diferencial para el anticuerpo anti-FLAG y anti-c-myc. **B)** Inmunofluorescencias de células COS-7 transfectadas con la construcción 4' que porta las secuencias codificantes para los epítopos FLAG y c-myc en los extremos 5' y 3', respectivamente. La tinción con el anticuerpo anti-FLAG se muestra en el panel de la izquierda, la del anticuerpo anti-c-myc en el panel central y la tinción nuclear en el panel de la derecha.

#### 1.8. La isoforma de 4.1R<sup>80</sup> generada *in vivo* desde el ATG2 no es un producto de proteólisis.

Para comprobar que en los experimentos *in vivo* la isoforma de menor tamaño estaba traducida desde el ATG2 y que no era un producto de proteólisis se transfectaron las construcciones 4' y 4 junto con los mutantes 4'STOP y 4'ATG2- en células COS-7. Los extractos de estas células se analizaron mediante *Western Blot* con los anticuerpos anti-c-myc y 762. En los extractos de células transfectadas con la construcción 4' el anticuerpo anti-myc detecta dos isoformas del tamaño esperado (fig. 22B izquierda) y de éstas sólo la de mayor peso molecular es reconocida por el anticuerpo anti-FLAG (fig. 22B derecha). En los extractos transfectados con la construcción 4'STOP, el anticuerpo anti-myc reconoce una única proteína que no es reconocida con el anticuerpo 762. Esta proteína presenta la misma movilidad electroforética que la proteína que genera la construcción 4 (comparar carriles 4' STOP y 4 fig. 22A).



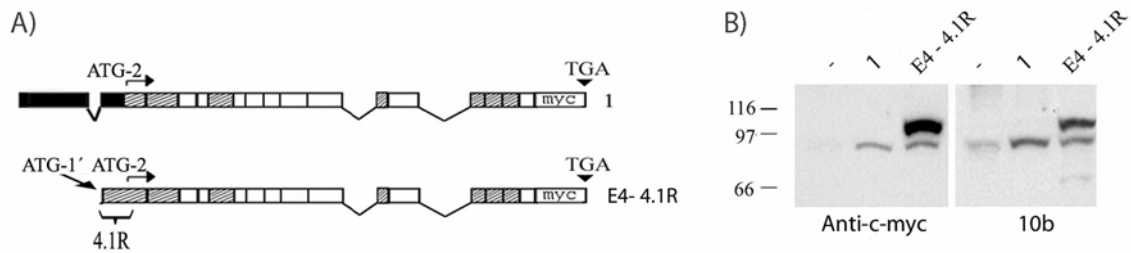
**Figura 22. La proteína de menor tamaño expresada desde el ADNc 4' transfectado en células COS-7 se produce a partir del ATG2 y no resulta de una proteólisis de la proteína de mayor tamaño. A)** Esquema de los ADNc utilizados en el apartado B. Los ADNc 4' y 4 codifican para isoformas naturales de 4.1R de 120 y 70 kDa respectivamente. En el ADNc 4' se sustituyó el codón de iniciación de la traducción ATG2 por el triplete GTG, esta construcción se denominó 4'ATG2-. En el ADNc 4' se introdujo un codón de parada de la traducción entre los codones de inicio de la traducción ATG1 y ATG2, esta construcción se denominó 4'STOP. **B)** Se transfectaron células COS-7 con los ADNc mostrados en el apartado A. Se analizaron mediante *Western Blot* los extractos proteicos totales recogidos a las 48h de la transfección con los anticuerpos anti-c-myc y 762. Encima de cada carril se indica el ADNc con el que se transfectaron las células COS-7 de cada extracto.

Por último, se detecta una sola proteína tanto con el anticuerpo anti-c-myc como con el anticuerpo 762 en los extractos de COS-7 transfectadas con la construcción 4'ATG2-. Los resultados obtenidos en este experimento indican que la proteína de menor tamaño es una isoforma de 4.1R y que, además, ésta se genera en ausencia de la isoforma de mayor peso molecular. Se descarta por lo tanto la implicación de un proceso proteolítico. Estos resultados son similares a los obtenidos en los experimentos de transcripción y traducción *in vitro*.

#### 1.9. La secuencia anterior al ATG2 del exón 4 de 4.1R genera la isoforma de peso molecular intermedio *in vivo*.

Extractos de células COS-7 transfectadas con las construcciones E4-4.1R (fig. 23A) y la construcción salvaje 1 se analizaron mediante *Western Blot* con los anticuerpos anti-c-myc y el anti-4.1R 10b (fig. 23B). La construcción E4-4.1R a pesar de no tener el exón 2, es capaz de producir dos proteínas. La menor de estas proteínas tiene una movilidad electroforética idéntica a la de la isoforma de 4.1R de 80 kDa (comparar carriles 1 y E4-4.1R), y además es reconocida no sólo por el anticuerpo anti-myc sino también por anticuerpos anti-4.1R (comparar carril E4-4.1R en *blot* anti-myc y 10b). Por lo tanto, la región 5' anterior al ATG2 de los ADNc para 4.1R<sup>135</sup> es capaz por sí misma, sin la presencia del exón 2, de dirigir *in vivo* la producción de la isoforma de 4.1R de peso molecular intermedio.



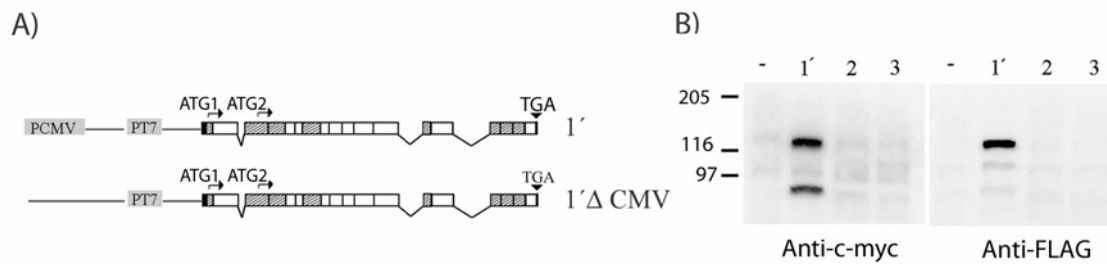


**Figura 23. La eliminación de los exones 2' y 2 permite la expresión de la isoforma de 4.1R de peso molecular intermedio.** A) Esquema de los ADNc utilizados en el apartado B. La primera construcción se corresponde con un ADNc para una isoforma natural de 4.1R<sup>80</sup> (1) y la segunda (E4-4.1R) es una construcción mutante que carece del exón 2 e introduce un codón de iniciación de la traducción (ATG1') en el extremo 5' del exón 4. B) Análisis mediante *Western Blot* de la expresión de los ADNc representados en el apartado A en células COS-7. El primer carril (-) muestra un control con el extracto proteico total de células COS-7 no transfectadas. En el resto de los carriles se ha indicado el ADNc con el que se transfectó a las células COS-7 cuyos extractos han sido analizados. El panel de la izquierda muestra el análisis con el anticuerpo anti-c-myc y el panel de la derecha el análisis con el anticuerpo anti-4.1R 10b.

#### 1.10. No existe un promotor críptico en el ADNc para 4.1R<sup>135</sup>.

La generación de dos isoformas a partir de un único ADNc puede estar dirigida por un promotor críptico que de lugar a un ARNm adicional al generado por el promotor presente en el vector que porta ese ADNc. Para descartar esta posibilidad se eliminó el promotor CMV del vector de expresión pCR3.1 en el que se había clonado el ADNc 1'. Al eliminar el promotor de este vector debería eliminarse la producción de cualquier proteína en sistemas eucariotas, salvo que exista un promotor interno capaz de dirigir la transcripción. Un esquema de esta construcción se muestra en la figura 24A. Se transfectaron en células COS-7 dos cantidades distintas (5 y 10 microgramos) de esta construcción, a la que hemos denominado 1' Δ CMV. Las transfecciones se realizaron por duplicado y se utilizó como control positivo la construcción 1' silvestre.

El análisis de estos extractos mediante *Western Blot* con los anticuerpos anti-c-myc y anti-FLAG no reveló la existencia de ninguna proteína (fig. 24B). En el extracto del control positivo, sin embargo, el anticuerpo anti-c-myc detecta la isoforma de 135 y de 80 kDa mientras que el anticuerpo anti-FLAG detecta únicamente la proteína de 135 kDa. Así, para que se produzcan las dos isoformas de 4.1R a partir de un único ADN que codifica para una isoforma de 4.1R<sup>135</sup> es necesaria la presencia del promotor CMV del vector pCR3.1. Por lo tanto no existe ningún promotor críptico dentro de la secuencia de 4.1 clonada en el vector pCR3.1.



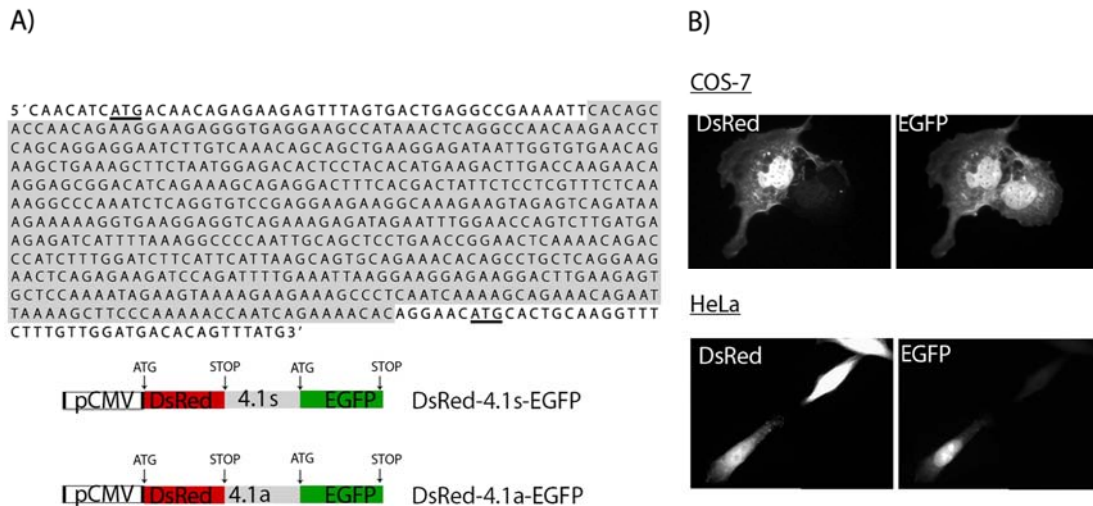
**Figura 24. No existe un promotor críptico en la secuencia de 4.1R que de lugar a un RNAm responsable de la producción de la isoforma de menor tamaño obtenida a partir de un ADNc para 4.1R<sup>135</sup>.** **A)** Esquema de los plásmidos utilizados en los *Western Blot* del apartado B. El primer plásmido (1') contiene el ADNc para una isoforma de 4.1R<sup>135</sup> y el promotor CMV. En el segundo plásmido se ha eliminado el promotor CMV (1'ΔCMV). Ambas isoformas contienen las secuencias codificantes para los epítomos FLAG y c-myc en los extremos 5' y 3' respectivamente. **B)** Análisis mediante *Western Blot* de los extractos de células COS-7 no transfectadas (-), transfectadas con el ADNc 1' (carril numerado 1'); transfectadas con 5 µg del plásmido 1'ΔCMV (carril 2) o con 10 µg del plásmido 1'ΔCMV (carril 3). El panel de la izquierda muestra el análisis con el anticuerpo anti-c-myc y el panel de la derecha el análisis con el anticuerpo anti-FLAG.

### 1.11. Generación de un vector bicistrónico.

Considerando todos los resultados obtenidos y la longitud de la secuencia comprendida entre los dos codones de iniciación ATG1 y ATG2 (627 nucleótidos) presentes en 4.1R, es razonable pensar que la síntesis de la isoforma de peso molecular intermedio esté dirigida por un proceso de traducción interna. Para comprobar la existencia de una traducción interna se construyó un vector bicistrónico. Este tipo de vectores están compuestos de dos cistrones que codifican para dos proteínas cuya síntesis puede determinarse cuantitativamente mediante diferentes técnicas. Entre ambos cistrones se clona la secuencia susceptible de dirigir una traducción interna. Si dicha secuencia puede dirigir una traducción interna se obtendrá la expresión de los dos cistrones. En nuestro caso el vector bicistrónico constaba de los cistrones que codifican para la proteína DsRed (proteína fluorescente del coral *Discosoma*) y la proteína EGFP (proteína verde fluorescente) en posición 5' y 3', respectivamente. Entre medias de ambos cistrones se clonaron 583 nucleótidos que abarcan la secuencia de 4.1 desde los 37 nucleótidos posteriores al ATG1 hasta los 6 nucleótidos anteriores al ATG2 (fig. 25A). Esta secuencia se clonó en orientación directa e inversa, utilizándose esta última como control negativo. Las construcciones se denominaron DsRed-4.1s-EGFP y DsRed-4.1a-EGFP, respectivamente.

### 1.12. Expresión del vector bicistrónico en células COS-7.

Las construcciones DsRed-4.1s-EGFP y DsRed-4.1a-EGFP se transfectaron en células COS-7 y la expresión de los cistrones DsRed y EGFP se analizó mediante inmunofluorescencia. La figura 25B muestra algunos ejemplos de los resultados obtenidos.



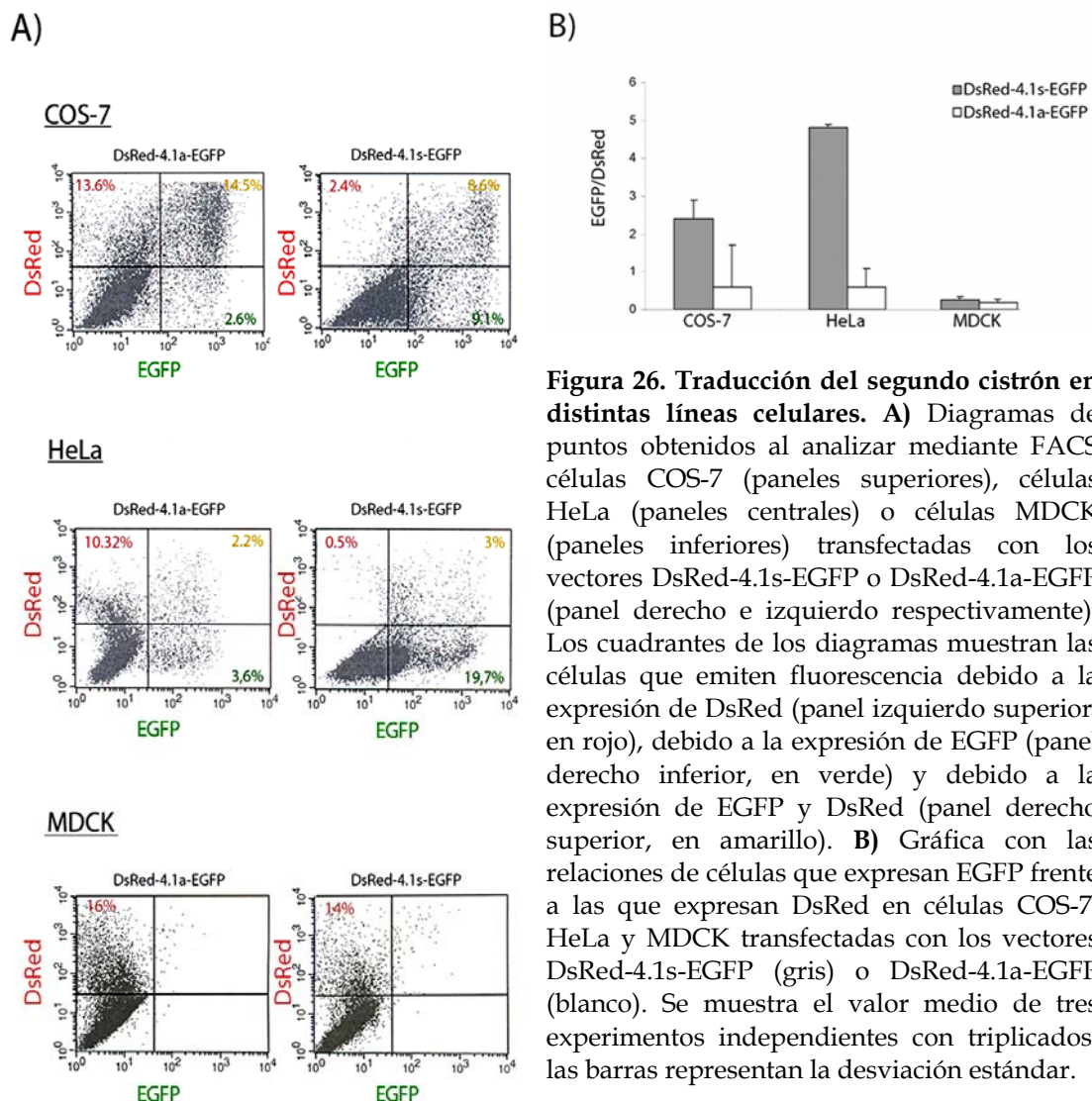
**Figura 25.** El test del vector bicistrónico indica que la secuencia de 4.1R comprendida entre el ATG1 y ATG2 es capaz de dirigir una traducción interna. **A)** En el panel superior se muestra la secuencia de 4.1R comprendida entre el ATG1 y ATG2 (ambos subrayados). En el panel inferior se muestra un esquema de los vectores bicistrónicos DsRed-4.1s-EGFP y DsRed-4.1a-EGFP. El gen que codifica para la proteína DsRed, situado después del promotor CMV y el gen que codifica para la proteína EGFP, situado después del gen de DsRed. La secuencia de 4.1R coloreada en gris en el panel superior se clonó entre ambos cistrones en orientación 5'-3' y 3'-5' en los vectores denominados DsRed-4.1s-EGFP y DsRed-4.1a-EGFP, respectivamente. **B)** Se fijaron células COS-7 (panel superior) y HeLa (panel inferior) transfectadas con los vectores bicistrónicos representados en el apartado A. Se analizaron mediante microscopía de fluorescencia. El panel de la izquierda muestra la fluorescencia emitida por la proteína DsRed y el de la derecha la emitida por la proteína EGFP.

En la línea celular COS-7 (paneles superiores) se observaron células que sólo tenían tinción roja, es decir, que sólo habían expresado el primer cistrón (DsRed), células que habían expresado únicamente el segundo cistrón, EGFP, y también células con expresión de ambos. En el caso de la línea celular HeLa, se obtuvieron resultados similares aunque la proporción de células que expresaban únicamente el primer cistrón parecía ser menor que en COS-7 (ver apartado siguiente). Sin embargo, la transfección de la línea celular MDCK con estas construcciones dio lugar tan solo a células que emitían en color rojo, es decir, que expresaban únicamente el primer cistrón (no se muestra). En todos los casos la construcción DsRed-4.1a-EGFP que portaba la secuencia de 4.1 en sentido reverso no produjo una cantidad

significativa de células que expresaran EGFP en comparación con la construcción DsRed-4.1s-EGFP, mientras que sí expresaron DsRed. Por lo tanto podemos concluir que el test del bicistrónico resultó positivo en las líneas celulares HeLa y COS-7.

### 1.13. Relación EGFP/DsRed en distintas líneas celulares.

Mediante citometría de flujo se cuantificó en las líneas celulares COS-7, HeLa y MDCK el número de células que expresaban únicamente el primer cistron, únicamente el segundo, o bien ambos.



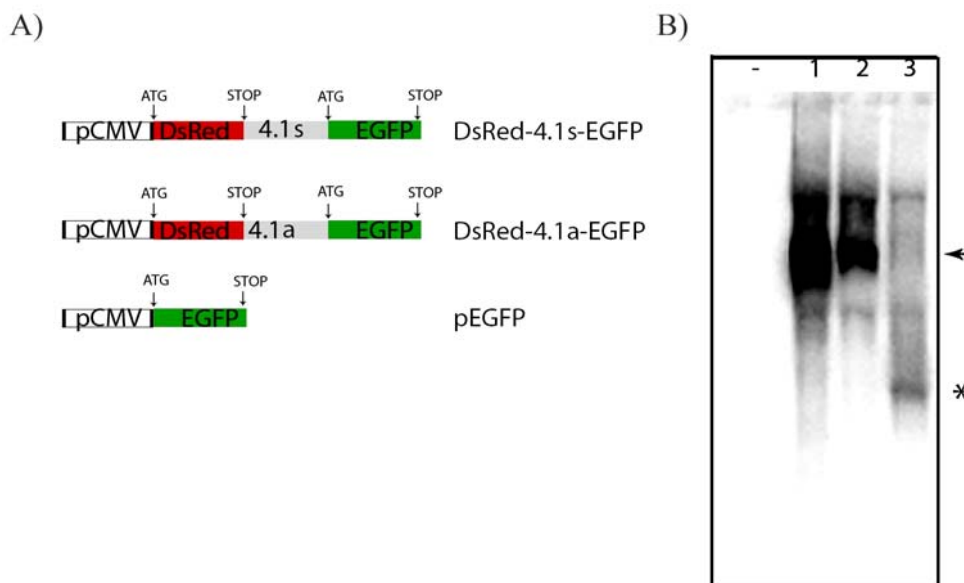
Los resultados obtenidos en uno de los experimentos representativos de citometría se muestran en los diagramas de puntos de la figura 26A. En el eje de abcisas de estos diagramas se registra la emisión de fluorescencia en verde debida a la expresión de EGFP y en el eje de ordenadas la emisión de fluorescencia roja debida a la expresión DsRed. Se indican, en los colores adecuados, los porcentajes de células que expresan DsRed, EGFP o ambos.

Como se puede observar el patrón de puntos que se obtiene para las células transfectadas con DsRed-4.1s-EGFP y las células transfectadas con DsRed-4.1a-EGFP es distinto tanto en células COS-7 como en HeLa; mientras que en células MDCK la distribución de las células en los distintos cuadrantes del diagrama de puntos es muy parecida.

La cuantificación de cinco experimentos con triplicados se muestra en el diagrama de barras de la figura 26B. En ella se representa la relación de células que expresan EGFP frente a aquellas que expresan DsRed. La relación de producción del segundo cistron (EGFP) frente al primero (DsRed) en células Hela es aproximadamente el doble que la de células COS-7. Asimismo, se puede observar que las células MDCK no presentan emisión de fluorescencia verde.

#### 1.14. La producción de EGFP no es consecuencia de un procesamiento del ARNm.

Se analizaron los ARNm producidos en células COS-7 transfectadas con DsRed-4.1s-EGFP mediante *Northern Blot*. Para ello se transfectaron células COS-7 con los plásmidos DsRed-4.1s-EGFP o DsRed-4.1a-EGFP o con el plásmido monocistrónico pEGFP control, que codifica únicamente para la proteína EGFP (fig. 27A).



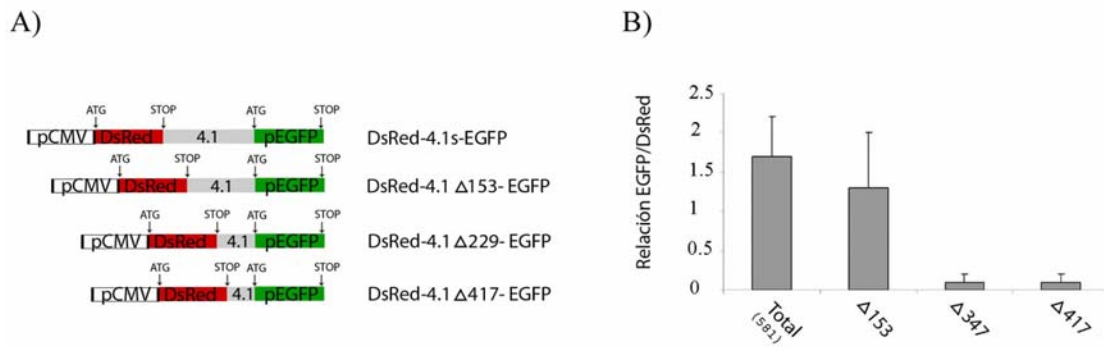
**Figura 27. Análisis de los ARNm que contienen el cistron que codifica para EGFP en células COS-7 transfectadas con el vector bicistrónico. A)** Esquema de las construcciones utilizadas en el apartado B. Se utilizaron los dos vectores bicistrónicos DsRed-4.1s-EGFP y DsRed-4.1a-EGFP así como un plásmido control que codifica para la proteína EGFP únicamente (pEGFP). **B)** Se extrajo el ARN total de células COS-7 no transfectadas (carril -) o transfectadas con los vectores representados en el apartado A; DsRed-4.1a-EGFP (carril 1), DsRed-4.1s-EGFP (carril 2) y pEGFP (carril 3) para analizarse mediante *Northern Blot* con una sonda cuya secuencia es la del ADNc para EGFP. Se indica mediante un asterisco la posición esperada para un ARN que codifica para EGFP y con una flecha la de los ARN monocistrónicos que codifican para DsRed y EGFP.



La sonda radiactiva utilizada para la hibridación comprende toda la secuencia del ADN codificante para EGFP. En la autorradiografía de la figura 27 (carril 1) se observa una banda del tamaño esperado para el ARN bicistrónico (2.120 pb) (se indica con una flecha). El control muestra el tamaño que se obtiene para un ARNm de EGFP (760 pb) (se indica con un asterisco). No se observa en ninguna de las construcciones bicistrónicas (DsRed-4.1s-EGFP y DsRed-4.1a-EGFP) un ARNm de este tamaño. Así, podemos concluir que existe un único ARNm responsable de la producción de DsRed y de EGFP.

#### **1.15. La secuencia de 4.1R clonada en el vector bicistrónico determina la eficiencia de producción del segundo cistron.**

Para determinar la secuencia de 4.1R necesaria para la producción de EGFP en células transfectadas con el vector bicistrónico, se realizaron tres mutantes de delección. En ellos se eliminaron 153, 229 y 417 nucleótidos del extremo 5'. Estas construcciones se denominaron DsRed-4.1 $\Delta$ 153-EGFP, DsRed-4.1 $\Delta$ 229-EGFP y DsRed-4.1 $\Delta$ 417-EGFP, respectivamente. Como control negativo para cada construcción, se clonó la misma secuencia en orientación inversa. Un esquema de estas construcciones se muestra en la figura 28A. Las diferentes construcciones se transfectaron de manera independiente y por triplicado en células COS-7. Se midió la expresión del segundo cistron con respecto al primero, cuantificando las células que emitían fluorescencia en verde y/o rojo mediante citometría de flujo. La relación de células que emitían fluorescencia en el canal del EGFP frente al número de células que emitían en el canal del DsRed se ha representado en la figura 28B. A esta relación se le restó siempre la relación EGFP/DsRed producida por la construcción control que contenía la secuencia de 4.1R clonada en sentido inverso. La secuencia que tenía una relación EGFP/DsRed más alta era aquella en la que no se había acortado secuencia del extremo 5' y que, por lo tanto, contenía 581 nucleótidos de 4.1R. La construcción que perdía 153 nucleótidos del extremo 5' presentaba también mayor expresión de EGFP que de DsRed pero la relación de éstos disminuye de 1,5, que presenta la construcción completa, a 1,3. Las otras dos construcciones que perdían 347 y 417 nucleótidos del extremo 5' de 4.1 presentaban una relación EGFP/DsRed muy parecida a la de los controles respectivos y, por lo tanto, no estaban produciendo de forma significativa el segundo cistron. Estos resultados indican que son necesarios al menos los 430 nucleótidos de la región 3' clonada en el vector bicistrónico, si bien la mayor eficiencia la tiene la construcción que contiene 583 nucleótidos de 4.1R.



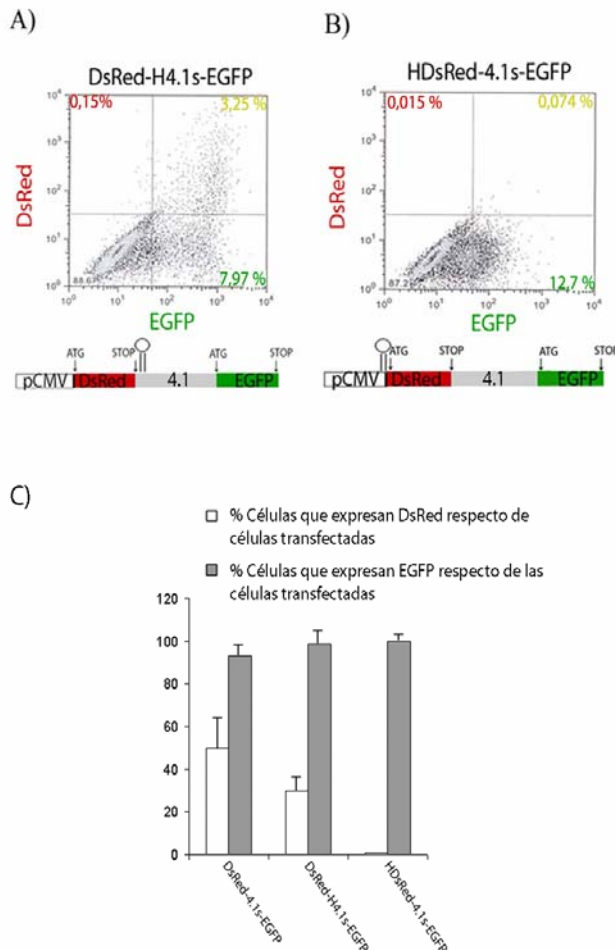
**Figura 28. La expresión del segundo cistron depende de la secuencia de 4.1R insertada en el vector bicistrónico.** A) Esquema de los vectores bicistrónicos utilizados en el apartado B. La secuencia de 4.1R clonada en estos vectores es de 583, 430, 354 o 166 nucleótidos (DsRed-4.1s-EGFP, DsRed-4.1Δ153-EGFP, DsRed-4.1Δ229-EGFP y DsRed-4.1Δ417-EGFP, respectivamente). El nombre de los vectores indica la secuencia perdida en el extremo 5' respecto de la secuencia de 581 nucleótidos de 4.1R clonada inicialmente. B) Gráfica de la relación de células COS-7 que expresan el segundo cistron (EGFP) frente al número de células que expresan el primero (DsRed). A esta relación se le restó siempre el valor de la relación obtenida para células transfectadas con vectores bicistrónicos controles en los que se clonó la secuencia de 4.1R en orientación 3'-5'. Las barras indican la desviación típica de los valores obtenidos en tres experimentos independientes.

#### 1.16. Los ADNc para 4.1R<sup>135</sup> con el exón 4 contienen una secuencia de entrada directa para el ribosoma (IRES).

Los resultados obtenidos hasta ahora sugieren que el mecanismo mediante el cual los ADNc para 4.1R<sup>135</sup> son capaces de generar una isoforma de 4.1R de peso molecular intermedio (80 kDa) está actuando a nivel de la traducción de éste. Existen diferentes mecanismos que permiten una traducción interna. Entre otras características, lo que diferencia a estos mecanismos es la necesidad o no de un rastreo del ribosoma por la zona anterior al codón de iniciación interno. Para poder discernir qué tipo de mecanismo está implicado, se clonó un fragmento de ADN que da lugar a un ARNm con una alta estructura secundaria, con un  $\Delta G^\circ$  de -57 Kcal/mol (Koromilas et al., 1992), que impide el deslizamiento del ribosoma. La construcción denominada DsRed-H4.1s-EGFP porta esta estructura detrás del primer cistron DsRed y antes de la secuencia de 4.1R (fig. 29A). El análisis, mediante citometría de flujo, de células COS-7 transfectadas con esta construcción no afectó a la expresión del cistron EGFP (fig. 29B), a pesar de que la estructura secundaria introducida impide el rastreo del ribosoma sobre el ARNm.

Se construyó otro vector bicistrónico que portaba la secuencia con alta estructura secundaria delante del primer cistron (fig. 29B). Esta construcción se denominó H-DsRed-4.1s-EGFP. Se analizaron mediante FACS células COS-7 transfectadas con esta construcción. No se observó ninguna célula capaz de expresar DsRed, como era de esperar. Sin embargo, la expresión de EGFP no se vio afectada (un 12.7% de células seguían expresando EGFP).

Los valores numéricos de las células que expresaban DsRed sobre el total de células transfectadas, así como de las que expresaban EGFP sobre las células transfectadas se representan en la gráfica 29C. Esta gráfica muestra que la expresión de EGFP no se ve afectada cuando se impide la expresión de DsRed.

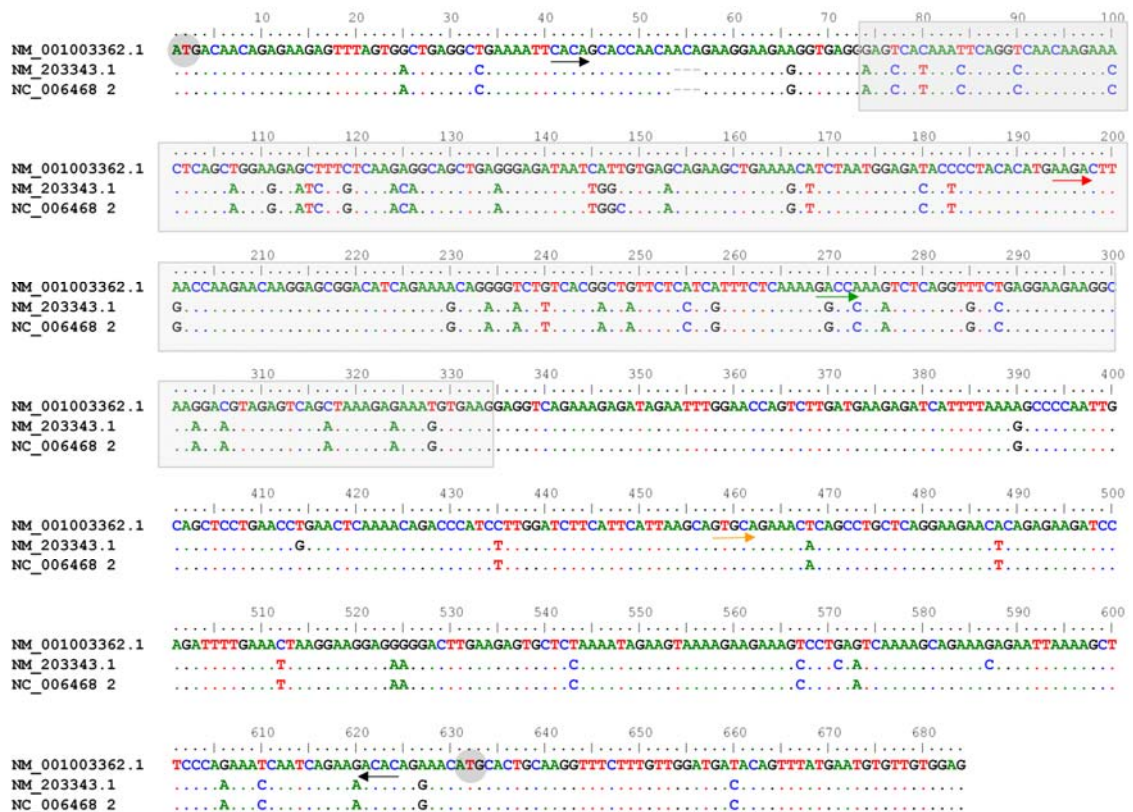


**Figura 29. El ribosoma se introduce directamente en la secuencia de 4.1R clonada en el vector bicistrónico.** A) Una secuencia con una alta estructura secundaria ( $\Pi$ ), ( $\Delta G^\circ$  -57 KCal/mol) se clonó en la zona 5' de la secuencia de 4.1R del vector bicistrónico DsRed-4.1s-EGFP. Esta construcción se denominó DsRed-H4.1s-EGFP y su esquema se muestra en la parte inferior del panel. En la parte superior se muestra el diagrama de puntos obtenido mediante FACS de células COS-7 transfectadas con esta construcción. B) La secuencia con alta estructura secundaria ( $\Pi$ ) se clonó en la zona 5' del primer cistón (DsRed) del vector bicistrónico. Esta construcción se denominó HDsRed-4.1s-EGFP y su esquema aparece en la parte inferior de este apartado. En la parte superior aparece el diagrama de puntos obtenido por FACS de células COS-7 transfectadas con esta construcción. C) Gráfica con los porcentajes de células COS-7 que expresan DsRed y células que expresan EGFP mediante FACS en células COS-7 transfectadas con DsRed-4.1s-EGFP, DsRed-H4.1s-EGFP o HDsRed-4.1s-EGFP. La estructura secundaria ( $\Pi$ ) no afecta a la expresión de EGFP.

### 1.17. Comparación de la secuencia comprendida entre el ATG1 y ATG2 del gen que codifica para 4.1R en diversas especies.

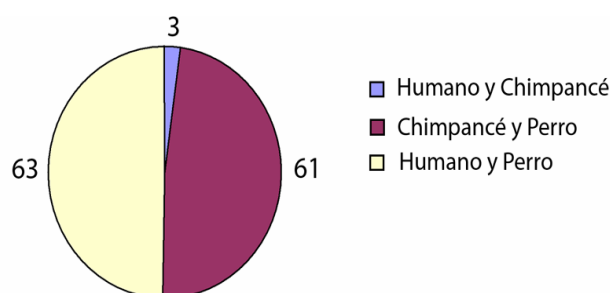
Dado que la expresión de EGFP a partir del vector bicistrónico que porta la secuencia de 4.1 no ocurre en células MDCK y sí en células HeLa y COS-7 realizamos una comparación de la secuencia comprendida entre el ATG1 y el ATG2 de 4.1R de humano, perro y chimpancé. Hay que puntualizar que utilizamos la secuencia de chimpancé, a pesar de que las células COS-7 proceden de mono verde africano, porque la secuencia de 4.1R de éste es desconocida.





**Figura 30. Comparación de la secuencia de 4.1R comprendida entre el ATG1 y el ATG2 de diferentes especies.** Se muestra la comparación de la secuencia de 4.1R comprendida entre el ATG1 y ATG2 (señalados con un círculo gris) de perro (NM\_001003362), humano (NM\_203343) y de chimpancé (NC\_006468). Los nucleótidos que coinciden se muestran con un punto y aquellos que faltan se muestran con un guión. En aquellos en los que existe un cambio se indica el nucleótido. Se ha señalado con flechas negras las secuencia de 4.1R humana clonada en el vector bicistrónico DsRed-4.1s-EGFP, la flecha roja, verde y naranja indica donde comienza la secuencia de 4.1R clonada en las construcciones DsRed-4.1  $\Delta$ 153-EGFP, DsRed-4.1  $\Delta$ 229-EGFP y DsRed-4.1  $\Delta$ 417 -EGFP, respectivamente. La secuencia que aparece sobre un fondo gris (261 nucleótidos) es aquella que presenta el mayor número de cambios.

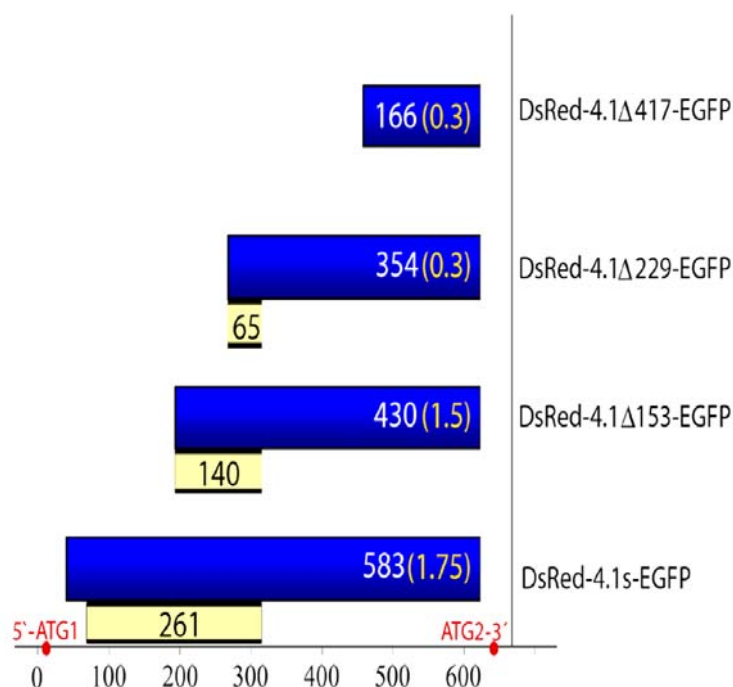
Como se puede observar la figura 30, y como era esperable, existen muchas más diferencias entre las secuencia de 4.1R de humano y perro que entre las de humano y chimpancé. En la gráfica de la figura 31 se han cuantificado estas diferencias. Existen sólo tres cambios entre las secuencias de humano



**Figura 31. Diferencias en la secuencia de 4.1R comprendida entre el ATG1 y el ATG2 de perro, humano y chimpancé.** La gráfica indica los nucleótidos que son diferentes en la secuencia de 4.1R comprendida entre el ATG1 y el ATG2 de humano y chimpancé, chimpancé y perro, y de humano y perro.

y chimpancé mientras que entre la secuencia humana y de perro existen 63. Asimismo, existen 3 nucleótidos compartidos en chimpancé y perro que cambian respecto de la secuencia humana. Estudiando la comparación de estas tres secuencias se puede observar que el fragmento que comprende del nucleótido 74 al nucleótido 335 presenta la mayor parte de los cambios. Esta región de 261 nucleótidos contiene 44 de los 63 cambios detectados en la secuencia humana y de perro (coloreadas en gris en la figura 30).

En el anillamiento de la figura 30 se han señalado mediante flechas las distintas secuencias de 4.1R que se han introducido en el vector bicistrónico. Como hemos descrito en el apartado anterior, las diferentes secuencias de 4.1R clonadas en el vector bicistrónico influían en la actividad que éste presentaba. En la figura 32 se muestra una gráfica en la que se puede observar que las construcciones que mantienen una producción significativa del segundo cistron del vector bicistrónico son aquellas que mantienen la mayor parte de la región de 261 nucleótidos donde se detecta el mayor número de cambios entre la secuencia de 4.1R humana y de perro. Asimismo se puede observar que una zona de esta región parece fundamental, ya que la eliminación de 116 nucleótidos del extremo 5' de la misma disminuyen la relación EGFP/DsRed de 1.75 a 1.5; mientras que la eliminación de otros 80 nucleótidos impide una producción significativa de EGFP.



**Figura 32. El acortamiento de la secuencia de 4.1R comprendida entre el ATG1 y ATG2 de humano afecta a la producción de EGFP.** En la gráfica se representa, mediante barras azules, la longitud de la secuencia de 4.1R (número en blanco) clonada en los vectores bicistrónicos indicados en el eje de ordenadas. El número en amarillo indica la relación EGFP/DsRed producida en células transfectadas con esos vectores bicistrónicos. El eje de abscisas indica los nucleótidos, el ATG1 y ATG2 de 4.1R. Las barras amarillas representan la longitud de la región que contiene el mayor número de cambios al comparar la secuencia humana y canina en cada una de las construcciones.

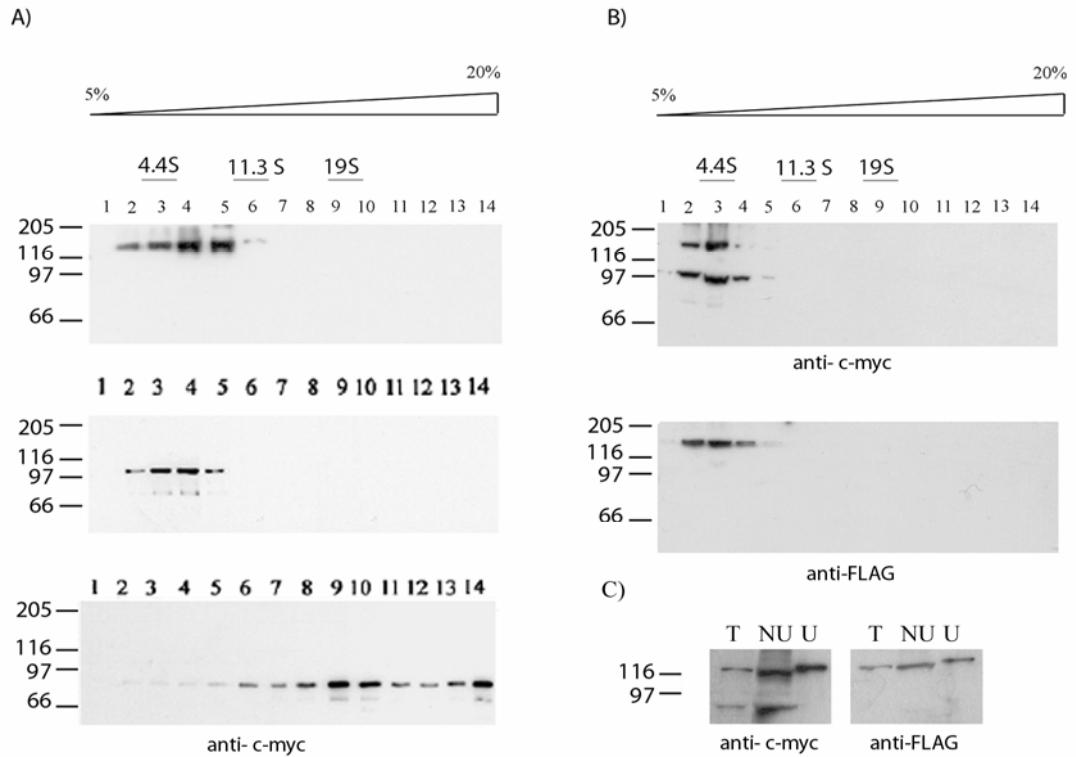
#### 1.18. Análisis de la traducción del segundo cistrón en células en cultivo sometidas a distintos tratamientos.

Se ha descrito la utilización de diferentes IRES en situaciones de estrés celular y también en algunas fases del ciclo celular en las que la traducción dependiente de la estructura CAP está inhibida. Se estudió si la relación EGFP/DsRed del vector bicistrónico variaba en estas condiciones. Se sometieron células COS-7 transfectadas transitoriamente con el vector bicistrónico a estrés térmico, osmótico, a la retirada de suero del medio, a la exposición de estaurosporina y de etopósido, así como a la sincronización con nocodazol con el fin de establecer si alguna de estas condiciones podía aumentar la relación del segundo cistrón frente al primero. En ningún caso se observó un aumento significativo de dicha relación en las células tratadas respecto a las células no tratadas.

#### 1.19. Las dos isoformas de 4.1R generadas desde un único ADNc no interactúan.

Nuestro grupo ha descrito que las isoformas de 4.1R de peso molecular bajo (60kDa) pueden autoasociarse formando complejos de coeficiente de sedimentación elevado. Las isoformas de peso molecular alto (4.1R<sup>135</sup>) e intermedio (4.1R<sup>80</sup>) no interactúan debido a que la región de oligomerización se encuentra inaccesible (fig. 33A). Una posible hipótesis para que las dos isoformas de 4.1R se generaran desde un único ADNc es que éstas pudieran interactuar antes de adoptar la conformación que ocultaría la zona de oligomerización. Para estudiar esta posibilidad se prepararon extractos de células COS-7 transfectadas con el ADNc 1' marcado con FLAG y c-myc en los extremos 5' y 3', respectivamente. Éstos se sometieron a centrifugación zonal en gradientes de sacarosa. El análisis mediante *Western Blot* de las distintas fracciones recogidas con los anticuerpos anti-FLAG y anti-c-myc se muestra en la figura 33A. Como era de esperar el anticuerpo anti-FLAG reconoce tan sólo a la isoforma de 4.1R de alto peso molecular mientras que el anticuerpo anti-c-myc detecta ambas isoformas. La distribución de estas dos isoformas es la misma a lo largo del gradiente y ninguna de ellas entra en la zona de mayor concentración de sacarosa. Estos resultados sugieren que las isoformas 4.1R<sup>135</sup> y 4.1R<sup>80</sup> generadas a partir de un mismo ADNc no se asocian formando complejos de alto coeficiente de sedimentación.

Ensayos de inmunoprecipitación de extractos de células COS-7 transfectadas con la construcción 1' con el anticuerpo anti-FLAG, analizados mediante *Western Blot* (fig. 33C), también indicaron que la isoforma de 4.1R de peso molecular alto no se asocia a la de peso molecular intermedio.



**Figura 33.** La isoforma 4.1R<sup>80</sup> expresada desde un ADNc para una isoforma para 4.1R<sup>135</sup> presenta el mismo comportamiento en gradientes de sacarosa que cuando es expresada desde un ADNc para 4.1R<sup>80</sup>. **A)** Se muestra el comportamiento de los tres tipos de isoformas de 4.1R en un gradiente de sacarosa cuando éstas son expresadas desde sus correspondientes ADNc, el panel superior corresponde a la isoforma de peso molecular alto (4.1R<sup>135</sup>) el central a una isoforma de peso molecular intermedio (4.1R<sup>80</sup>) y el inferior a la isoforma de peso molecular bajo (4.1R<sup>135</sup>) que es la que se asocia y por lo tanto aparece en las fracciones de mayor densidad. **B)** Se muestra el comportamiento de las dos isoformas de 4.1R expresadas desde el ADNc 1' en gradientes de sacarosa analizados mediante *Western Blot* con el anticuerpo anti-c-myc (panel superior) y con el anticuerpo anti-FLAG (panel inferior). La isoforma de 4.1R<sup>80</sup> expresada desde este ADNc se comporta de igual forma en un gradiente de sacarosa que si se hubiera expresado desde un ADN para la isoforma de 4.1R<sup>80</sup>. **C)** Inmunoprecipitación de células COS-7 transfectadas con la construcción 1' marcada con los epítomos FLAG y c-myc en 5' y 3' respectivamente analizadas mediante *Western blot* con los anticuerpos indicados. (T: total, UN: no unido, N: unido).

## 2. AISLAMIENTO DE NUEVOS ADNc PARA ISOFORMAS DE 4.1R QUE CARECEN DEL DOMINIO CARBOXILO CARACTERÍSTICO (CTD).

Diversas instituciones han desarrollado programas que hacen accesibles las secuencias validadas y los clones de diferentes ADN codificantes que contengan la información completa para codificar una proteína. Estos ADNc son extremadamente útiles para determinar la estructura de los genes y la forma en que éstos se están expresando en un determinado tejido o célula. Durante el desarrollo de este trabajo uno de estos proyectos MGC (*Mammalian Gene Collection*) del NIH (*The National Institutes of Health*) generó e identificó 15.000 ADN codificantes completos de humanos y de ratón. Uno de estos ADN codificantes completos, obtenido de la librería NIH\_MGC\_119 de cerebro humano (Life Technologies, Inc.), contiene la secuencia que codifica para una isoforma de 4.1R que no se había descrito con anterioridad.

Este ADNc, con número de acceso en el genbank BC039079, centró nuestra atención por presentar un tipo de procesamiento alternativo en el exón 17 no descrito anteriormente. Se propuso como objetivo de este trabajo estudiar este ADNc y la isoforma que genera, así como la posible existencia de más ADNc para 4.1R con este procesamiento.

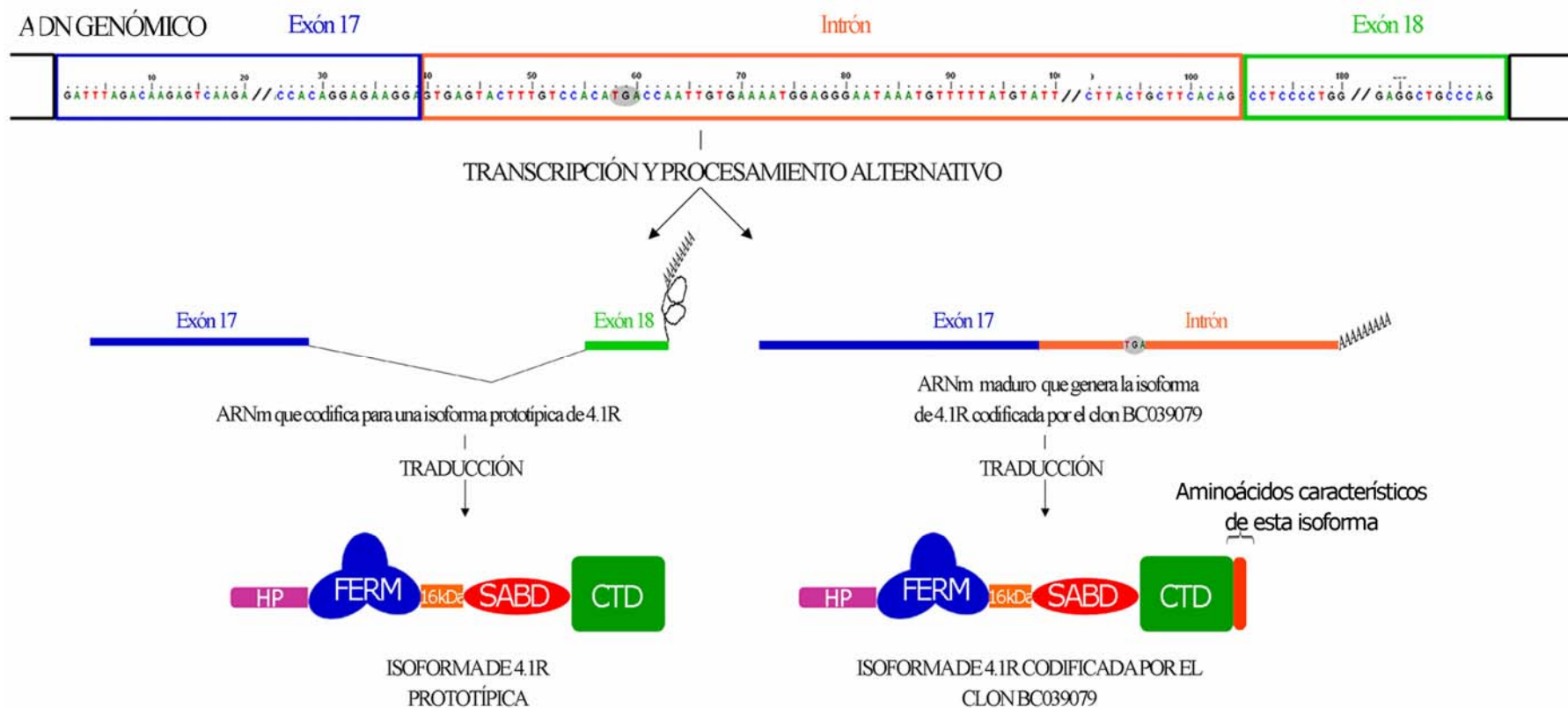
### 2.1. Características del clon BC039079.

Los ADNc para 4.1R caracterizados hasta el año 2002 terminan su traducción en un codón de parada localizado en el exón 21. A éste le sigue el exón 22 de más de 3 Kb que no es traducido y constituye la región 5' UTR de los ARNm de 4.1R (Baklouti 1997). El clon BC039079, sin embargo, presenta la particularidad de terminar en el exón 17, en el que además se mantiene parte de la secuencia del intrón (61pb) localizado entre los exones 17 y 18 del ADN genómico (fig. 34). La secuencia del intrón que se mantiene en este clon, que no poseen el resto de ADNc aislados hasta la fecha, introduce una secuencia que codifica para seis aminoácidos característicos de esta isoforma, detrás de la cual hay un codón de parada de la traducción y una señal de poli-A. La traducción de este ARNm genera una isoforma de 4.1R que carece de la zona carboxilo terminal correspondiente al dominio carboxilo de 22-24 kDa de 4.1R (CTD). Además de perder esta zona, la isoforma de 4.1R sin dominio CTD presenta los aminoácidos VSTLST, que no poseen el resto de isoformas de 4.1R.

### 2.2. Expresión del clon BC039079.

Se adquirió el clon BC039079 para estudiar su expresión. Este clon contiene los exones 2', 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16 y 17 (fig. 36A). Se analizaron extractos de células COS-7 transfectadas con él mediante *Western Blot* con los anticuerpos anti-4.1R: 10a y 10b (fig. 35A), que reconocen epítomos en los exones 16 y 17, respectivamente.





**Figura 34. Esquema del procesamiento alternativo sufrido por el ARNm que da lugar al ADNc del clon BC039079 y de la proteína que éste genera en comparación con el de una isoforma de 4.1R típica.** La parte superior del esquema muestra el fragmento del ADN genómico que codifica para las isoformas de 4.1R comprendido entre los exones 17 y 18 (en azul y verde respectivamente). La transcripción, procesamiento y maduración del ARNm que da lugar al ADNc que contiene el clon BC039079 se muestra en el panel de la derecha. Este ARNm maduro termina antes que el resto de los ARNm de 4.1R debido a un procesamiento alternativo y a una señal de poliadenilación prematura que producen la incorporación de un fragmento del intrón 17. La traducción de este ARNm genera una isoforma de 4.1R que carece del dominio carboxilo CTD que presentan el resto de isoformas de esta proteína. Se ha señalado el codón de parada de la traducción con un círculo en gris. En el panel de la izquierda se representa el ARNm maduro de una 4.1R prototípica. La consecuencia final es que el clon BC039079 codifica para una isoforma de 4.1R que carece del dominio carboxilo característico del resto de isoformas de 4.1R.

El número de aminoácidos esperados para la proteína codificada por el clon BC039079 es de 726 y el peso molecular estimado es de unos 82 kDa aunque es conocido que la mayor parte de isoformas grandes de 4.1R presentan una movilidad electroforética menor de la esperada. Como se puede observar en la figura 35A los dos anticuerpos anti-4.1R reconocen una única proteína de 105 kDa. Nos referiremos a partir de ahora a esta isoforma que carece del dominio CTD, como clon o proteína de referencia.

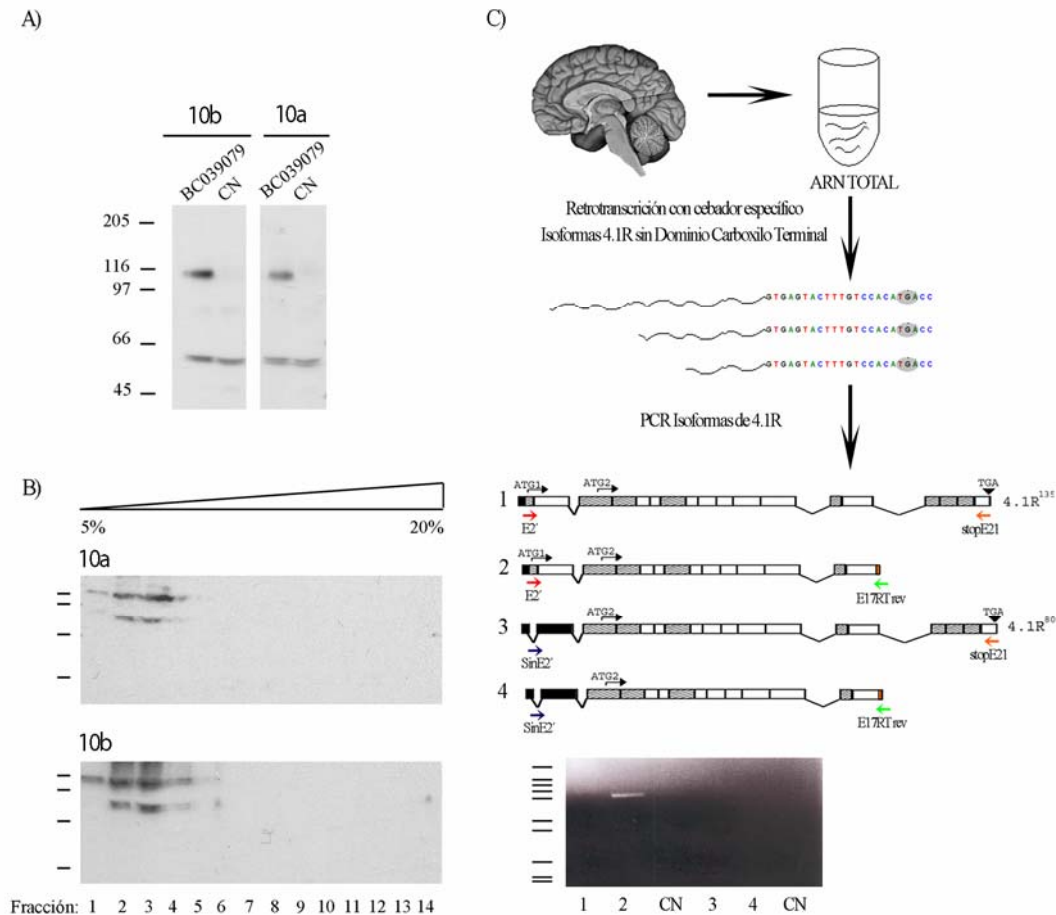
Se analizó el comportamiento de esta nueva isoforma de 4.1R en gradientes de sacarosa para determinar si pertenecía al grupo de isoformas de 4.1R capaces de autoasociarse. En la figura 35B se muestra que esta isoforma se aísla en las fracciones del gradiente de bajo coeficiente de sedimentación. Por lo tanto podemos concluir que esta isoforma no se autoasocia.

### 2.3. Nuevas isoformas sin dominio CTD en cerebro humano.

Se planteó a continuación determinar si existían otras isoformas de 4.1R que careciesen del dominio CTD. Para ello se realizó una retrotranscripción de ARN de cerebro humano (fig. 35C). Se empleó un cebador específico para la zona 3' del exón 17 características de la secuencia de ADN que codifica para la proteína de referencia. Seguidamente se realizó una amplificación de los ADNs obtenidos en la reacción de retrotranscripción, mediante la reacción en cadena de la polimerasa (PCR). Se utilizó para ello cebadores específicos para isoformas que tuvieran el exón 2' (cebador E2' y 17 RT rev). En todas las reacciones se realizaron dos controles negativos uno que no contenía cebadores (CN) y otro (carril 1 y 4) en el que se cambiaba el cebador del extremo 3' (E17RT rev) por un cebador específico para las 4.1R con dominio CTD (E20A). Se intentó amplificar también ADNs que codificaran para isoformas de 4.1R que carecen del exón 2' (cebador sinE2' y 17 RT rev) (fig. 35C).

Se obtuvo una reacción positiva con la pareja de cebadores E2'S y E17RTrev, que dio lugar a un ADN de 2000 pb como se puede observar en la figura 35 (carril 2). Aunque se utilizaron diferentes condiciones de amplificación y distintos cebadores para amplificar ADNc sin dominio CTD que carecieran del exón 2', nunca se obtuvo una reacción de amplificación positiva ni en el ARN de cerebro humano (carril 4) ni en el ARN de las líneas celulares SH5YSY y L6-IC (no se muestra).

El ADN de la reacción de amplificación positiva fue clonado en el vector pcDNA3.1/CTGFP-topo. Se seleccionaron clones positivos mediante corte con enzimas de restricción. En total se obtuvieron 11 clones positivos, siendo todos ellos secuenciados.



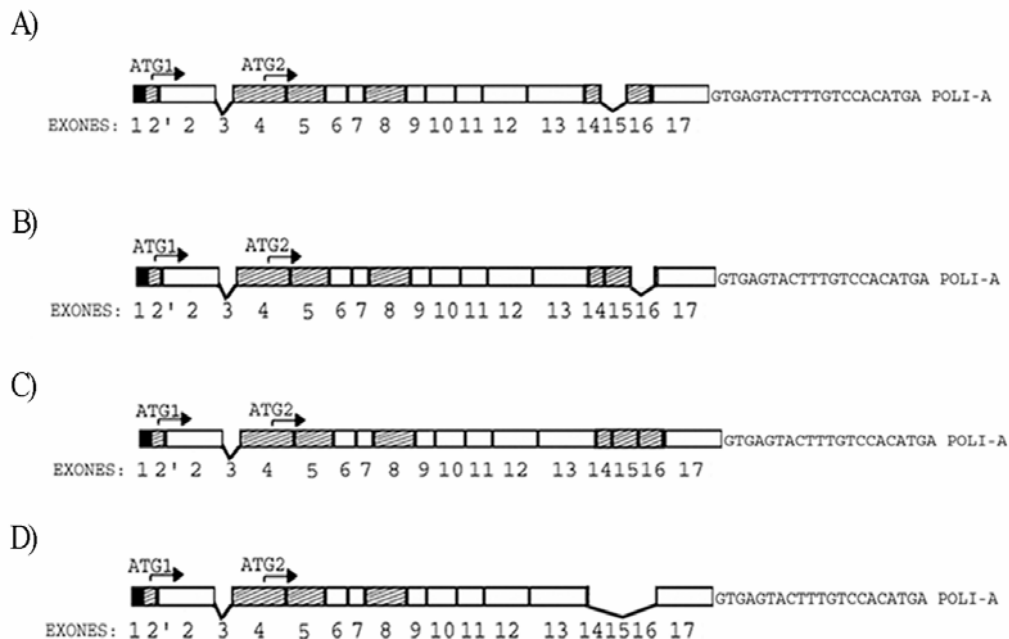
**Figura 35. Expresión del clon BC039079 en células COS-7 y aislamiento de nuevos ADNc para isoformas de 4.1R sin CTD.** **A)** Se analizaron extractos de células COS-7 transfectadas con el vector que contenía el ADNc del clon BC039079, mediante *Western Blot* con los anticuerpos anti-4.1R 10b y 10a. El carril denominado CN (control negativo) contiene un extracto de células COS-7 no transfectadas. **B)** Se analizó el comportamiento de la isoforma de 4.1R sin CTD en gradientes de sacarosa, cuyas fracciones se analizaron mediante *Western Blot* con los anticuerpos anti-4.1R 10a y 10b. **C)** Un extracto de cerebro humano se retrotranscribió con un cebador específico de la secuencia codificante para la isoforma de 4.1R sin CTD. Se amplificaron ARNm con cebadores específicos para isoformas de 4.1R de peso molecular alto con y sin CTD (carriles 1 y 2 respectivamente) y para isoformas de peso molecular intermedio con y sin CTD (carriles 3 y 4 respectivamente). Los cebadores para isoformas con CTD se utilizaron como control. Los carriles nombrados como CN corresponden a controles negativos en los que no se añadió el molde de ADN.

#### 2.4. Composición exónica de los ADNc aislados.

De los 11 clones aislados, 3 presentaban una composición exónica como la del clon BC03979. Los 8 clones restantes se clasificaron en tres grupos según su composición exónica. En la figura 36 se muestra un esquema de la composición exónica de cada uno de estos grupos. Cuatro clones se diferenciaban del clon de referencia por la falta del exón 16 y la presencia del exón 15 (clones: 8, 33, 40 y 45). El siguiente grupo se caracterizaba por carecer del exón 14, exón que mantiene el clon de referencia, además de los



exones 15 y 16 (clon 5 y 29). El último grupo lo constituye un único clon, el 32, que se caracteriza por tener tanto el exón 14 como el 15 y el 16.



**Figura 36. Clasificación de los ADNc para isoformas de 4.1R sin CTD en función de su composición exónica.** Se muestra el esquema de los ADNc aislados de cerebro humano. Los exones se han representado como sigue: alternativos, en gris; constitutivos, en blanco; no codificantes, en negro. Se ha indicado la secuencia nucleotídica 3' específica de estas isoformas. El número de cada exón se indica en la parte inferior de los esquemas. Se indican los dos codones de iniciación de la traducción, situados en los exones 2' y 4 (ATG1 y ATG2 respectivamente). **A)** Esquema de los ADNc para las isoformas denominadas 3, 4, 6 y 19, así como la del clon BC039079. **B)** Esquema de los ADNc para las isoformas denominadas 8, 33, 40 y 45. **C)** Esquema del ADNc para la isoforma 32. **D)** Esquema de los ADNc para las isoformas denominadas 5 y 29.

La diferencia más destacable es que dos de estos clones (33 y 29) presentan una delección de tres nucleótidos en la posición 1641-1643 que elimina un triplete y por lo tanto un aminoácido en el exón 13. En la tabla 3 se muestran los pesos moleculares estimados de cada una de las proteínas que generarían los clones que hemos aislado, el número de aminoácidos que tienen, su punto isoeléctrico y el número de residuos cargados positiva y negativamente.

Isoforma	Nº aa	Pm (Da) estimado	PI estimado	Residuos con cargas negativas (Asp+Glu)	Residuos con cargas positivas (Arg+Lys)	E14	E15	E16
Clon 3	720	81.235	5,57	123	103	SI	NO	SI
Clon 4	720	81.235	5,57	123	103	SI	NO	SI
Clon 6	720	81.235	5,57	123	103	SI	NO	SI
Clon 19	720	81.235	5,57	123	103	SI	NO	SI
Clon 8	713	80.325	5,58	122	103	SI	SI	NO
Clon 33	712	80.250	5,58	122	103	SI	SI	NO
Clon 40	713	80.264	5,58	122	103	SI	SI	NO
Clon 45	713	80.325	5,58	122	103	SI	SI	NO
Clon 5	680	76.580	5,62	115	98	NO	NO	NO
Clon 29	679	76.485	5,53	116	97	NO	NO	NO
Clon 32	734	82.922	5,66	126	108	SI	SI	SI

**Tabla 3.** En la tabla se indican las características de las proteínas codificadas por los ADNc denominados 3, 4, 6, 19, 8, 33, 40, 45, 5, 29 y 32, el número de aminoácidos de éstas, el peso molecular estimado, el punto isoelectrico y los residuos cargados negativa y positivamente, así como la presencia de los exones 14, 15 y 16.

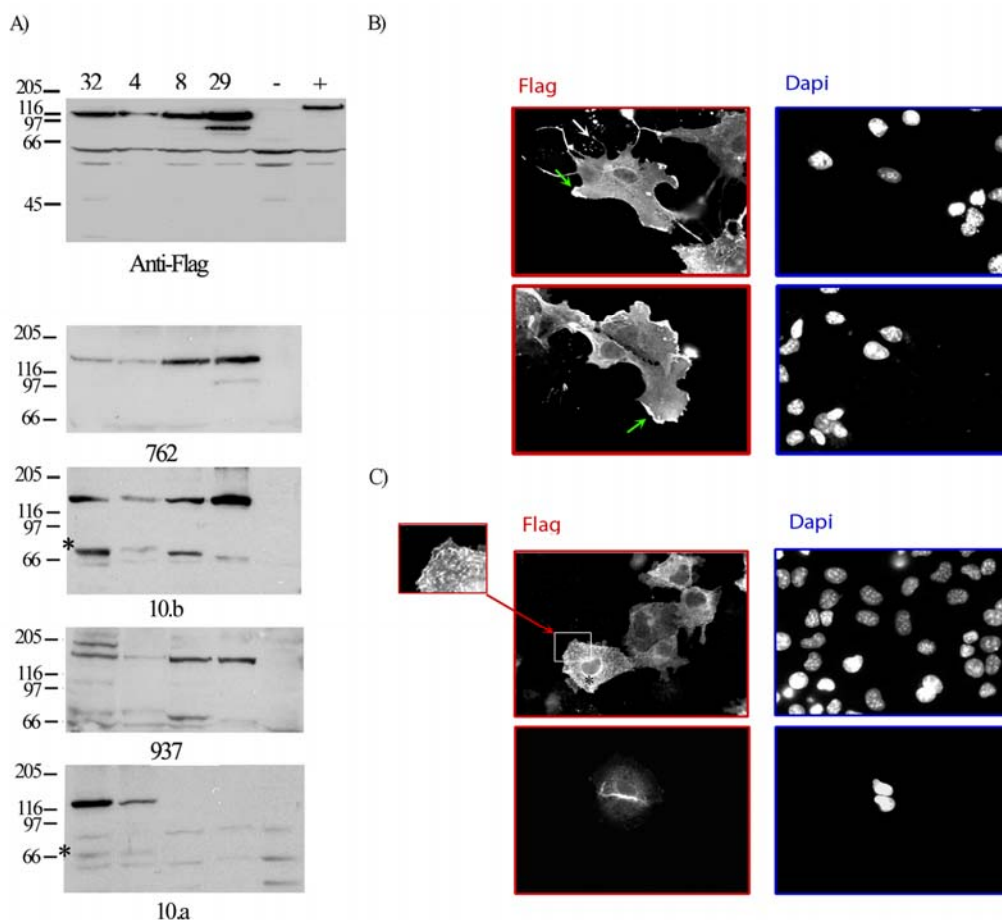
## 2.5. Expresión de los ADNc para isoformas de 4.1R sin dominio CTD.

Se procedió a continuación a subclonar en el vector de expresión eucariota pCR3 una isoforma representante de cada grupo. Elegimos los clones 4, 8, 29 y 32. Cada uno de ellos se marcó en el extremo 5' con el epítipo FLAG. Este epítipo permitiría identificar específicamente las proteínas generadas desde estas construcciones sin confundirlas con las isoformas endógenas de 4.1R. Tras su clonación se secuenciaron nuevamente los cuatro clones, comprobando que no se habían producido cambios en la secuencia. Se analizaron extractos de células COS-7 transfectadas con estos ADNc mediante *Western Blot* con distintos anticuerpos.

El análisis de estos extractos con el anticuerpo anti-FLAG se muestra en la figura 37A. Podemos observar que el anticuerpo anti-FLAG reconoce a las cuatro isoformas subclonadas en el vector pCR3 y que sus movilidades electroforéticas son muy similares, ya que son muy pocos los aminoácidos que las diferencian. La mayor es la isoforma 32, a la que siguen en tamaño decreciente las isoformas 4, 8 y 29. El tamaño aproximado que presentan es de unos 105 kDa. Este tamaño es en todos los casos mayor que el estimado (ver tabla 3). Como cabía esperar, se observa una diferencia apreciable con respecto al tamaño de una isoforma de 4.1R que tiene dominio CTD (carril + fig. 37A).

Se analizaron los mismos extractos con diferentes anticuerpos anti-4.1R existentes en el laboratorio (fig. 37A). Los anticuerpos utilizados fueron: 762, 10a y 10b que reconocen epítipos en los exones 2, 16 y 17 respectivamente. Se utilizó también el anticuerpo 937 que se obtuvo en el laboratorio

utilizando como antígeno la zona de 4.1R que comprende la secuencia codificada por los exones 17 al 20. Todos los anticuerpos reconocieron una proteína del tamaño esperado para las construcciones transfectadas. Como cabía esperar el anticuerpo 10a no reconoce a las proteínas codificadas por los ADNc 8 y 29 ya que éstas no expresan el exón 16. El anticuerpo 937 reconoce todas las isoformas ya que todas poseen el exón 17. Los anticuerpos 10a, 10b y 937 además de detectar la proteína del tamaño esperado detectan también una proteína (no presente en los controles negativos) de menor tamaño (fig. 37A asteriscos). Teniendo en cuenta los resultados de apartados anteriores, estas bandas podrían corresponder a isoformas traducidas desde el ATG2 de estos ADNc.



**Figura 37. Expresión de los ADNc 4, 8, 29 y 32 en células COS-7. A) Análisis mediante Western Blot.** Extractos de células COS-7 no transfectadas (-) o transfectadas con los ADNc 32, 4, 8 y 29 o con un control positivo (ADNc para una isoforma de 4.1R con el dominio CTD) se analizaron mediante Western Blot con el anticuerpo anti-FLAG y con diferentes anticuerpos anti-4.1R (762, 10b, 937, 10a). **B) y C) Análisis mediante inmunofluorescencias.** Se transfectaron los ADNc para las isoformas de 4.1R sin CTD (29 y 4, respectivamente) en células COS-7 y se analizaron mediante inmunofluorescencias con el anticuerpo anti-FLAG (paneles de la izquierda). Se muestra la tinción nuclear con DAPI (paneles de la derecha). Se ha señalado con flechas algunas localizaciones en las que se acumulan estas isoformas (flechas verdes en la membrana plasmática, flecha blanca en las prolongaciones de esta). En el apartado C se amplía una zona de la membrana plasmática en la que se observan ondulaciones en las que se acumulan estas isoformas.

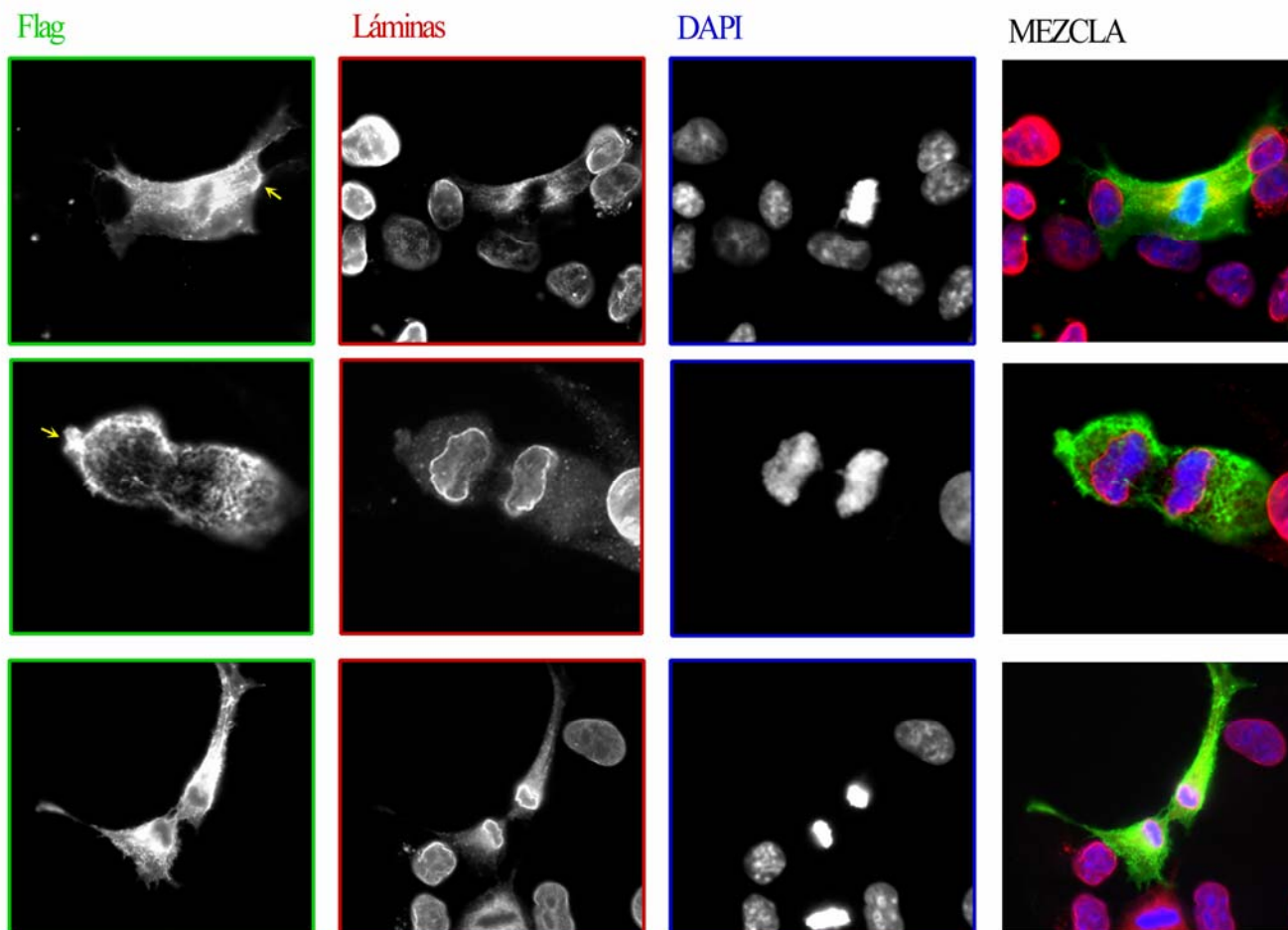
## 2.6. Localización intracelular de las isoformas de 4.1R sin dominio CTD.

Se transfectaron células COS-7 con los cuatro ADNc correspondientes a los clones 4, 8, 29 y 32 y se fijaron 24 h después de la transfección. Se analizó la localización de la proteína expresada mediante inmunofluorescencia con el anticuerpo anti-FLAG. La forma de las células transfectadas con los diferentes ADNc para isoformas sin dominio CTD es diferente de la que presentaban las células no transfectadas. Presentaban un tamaño mayor que las células control y emitían prolongaciones de la membrana plasmática (fig. 37B y 37C). También hay que destacar que aparecen muchas células binucleadas (no se muestra). En cuanto a su localización, se distribuyen en el citoplasma de la célula originando una tinción muy marcada en determinadas zonas de la membrana plasmática como en ondulaciones, lamelipodios, filopodios (fig. 37B y 37C), en los contactos entre algunas células (fig. 37C) y en algunas zonas de la membrana nuclear. No se observan diferencias muy acusadas entre los distintos ADNc transfectados. En general la transfección del ADNc 29 da lugar a células con unas prolongaciones muy largas que dan un aspecto de “neurona” a la célula, mientras que las células transfectadas con el ADNc 4 presentan un alto número de ondulaciones de membrana.

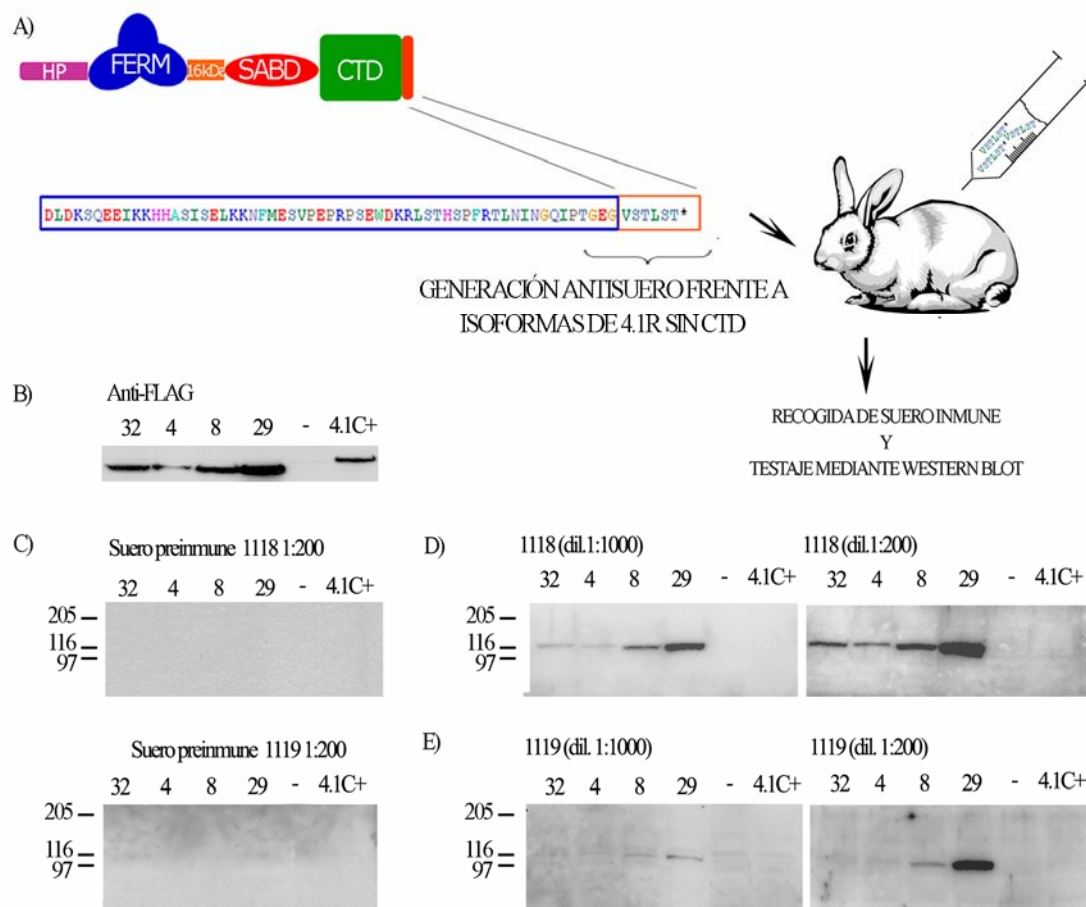
Con el fin de seguir la distribución de las isoformas de 4.1R sin dominio CTD a lo largo del ciclo celular, se realizaron tinciones dobles de células COS-7 transfectadas con los diferentes ADNc con los anticuerpos anti-FLAG y anti-láminas. Como se puede observar en la figura 38 en aquellas células que están en metafase, las isoformas de 4.1R sin dominio CTD se concentran en la zona que rodea a la placa metafásica. Desde ese momento hasta que las células hijas se han separado, la localización de las isoformas sin CTD es menos homogénea que el de células interfásicas, presentando un aspecto mas filamentoso y concentrándose en la zona que rodea al ADN (fig. 38).

## 2.7. Generación de un anticuerpo específico para las isoformas de 4.1R sin dominio CTD.

Como ya se ha comentado, las isoformas que se han aislado en este trabajo se caracterizan no sólo por la falta del extremo CTD sino que además tienen en su extremo carboxilo seis aminoácidos específicos (VSTLST) que no presentan las isoformas con CTD. Utilizamos un péptido (KTGEGVSTLST) que contenía, entre otros, estos seis aminoácidos (VSTLST) para generar un anticuerpo policlonal específico (fig. 39A). Los sueros preinmune e inmune se testaron mediante *Western Blot* en extractos de células COS-7 transfectadas con los diferentes ADNc para las isoformas de 4.1R sin CTD. Como control negativo se utilizó un extracto de células COS-7 transfectadas con una isoforma de 4.1R que mantiene el dominio CTD. La expresión de las diferentes isoformas se ratificó mediante con el anticuerpo anti-FLAG (fig. 39B) En estos extractos, los dos antisueros (1118 y 1119) reconocieron específicamente las isoformas de 4.1R sin CTD y no detectaron a la isoforma de 4.1R con dominio CTD.



**Figura 38. Expresión de las isoformas de 4.1R sin dominio CTD durante la división celular.** Inmunofluorescencias de células COS-7 transfectadas con el ADNc 4. Se muestran fotos representativas de la tinción obtenida para los anticuerpos anti-FLAG y anti-láminas así como la tinción con DAPI y la mezcla obtenida en todos los casos. Se observa un acúmulo de estas isoformas alrededor del ADN condensado y también en algunas zonas de la membrana (señalado con una flecha amarilla).



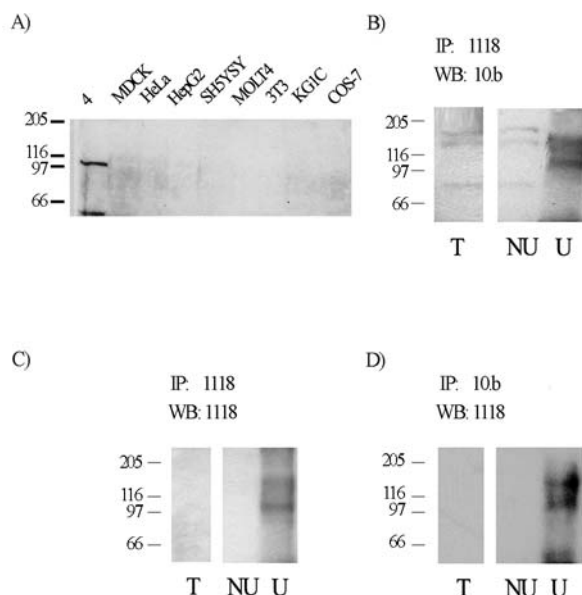
**Figura 39. Generación de un antisuero específico para las isoformas de 4.1R sin dominio CTD.** A) Se muestra la composición aminoacídica del exón 17 (cuadro azul) que tienen todas las isoformas de 4.1R así como la secuencia específica de las isoformas sin CTD (recuadro naranja). Se utilizó la secuencia señalada con una llave para generar un péptido que se pinchó en dos conejos para generar un suero específico contra las isoformas de 4.1R sin CTD. B) Western Blot de extractos de células COS-7 no transfectadas (-) o transfectadas con los ADNc 32, 4, 8, 29 o con un ADNc usado como control positivo que corresponde a una isoforma de 4.1R con dominio CTD (4.1C+). C, D) y E) Los mismos extractos de células COS-7 utilizados en el apartado B se analizaron mediante Western Blot con los suero preinmunes de los dos conejos utilizados para generar el antisuero (C), o con el suero preimmune de uno de los conejos, 1118 (D), o el antisuero del segundo conejo inmunizado, 1119 (E).

## 2.8. Análisis de la expresión endógena de las isoformas 4.1R sin CTD en distintos tipos celulares.

Se obtuvieron extractos totales de diferentes líneas celulares (MDCK, HeLa, HepG2, SH5YSY, MOLT4, 3T3, KG1C y COS-7) con el objetivo de identificar aquellas que expresasen de forma endógena las isoformas de 4.1R sin CTD. Estos extractos se analizaron mediante Western Blot con el antisuero 1118. Como control positivo se utilizó un extracto de células COS-7 transfectadas con la isoforma 4 (carril 4 fig. 40). En ninguno de los extractos totales se detectaron isoformas de 4.1R sin CTD. Sin embargo, el anticuerpo 1118 sí



detectó estas isoformas en ensayos de inmunoprecipitación de extractos totales de células COS-7, HepG2, y HeLa. En todas estas líneas celulares se detectaron dos isoformas mediante *Western Blot* con el anticuerpo 1118 cuando se inmunoprecipitaba un extracto total de estas con ese mismo suero. Se obtenían proteínas con idéntica movilidad electroforética cuando se inmunoprecipitaba con el antisero 1118 y se analizaba la inmunoprecipitación con el anticuerpo 10b que reconoce a todas las isoformas de 4.1R (tengan o no dominio CTD); o cuando se inmunoprecipitaba con 10b y se analizaba con el antisero 1118. De las dos isoformas reconocidas en todas las líneas una de ellas presenta una movilidad electroforética comprendida entre 130 y 97 kDa y otra comprendida entre 97 y 80 kDa, dependiendo de la línea celular. En la figura 40 se muestran de forma representativa los resultados obtenidos en la línea celular COS-7.



**Figura 40. Isoformas de 4.1R sin dominio CTD en diferentes líneas celulares.** **A)** *Western Blot* de extractos totales de las diferentes líneas celulares indicadas (parte superior de los carriles) con el antisero 1118. El control positivo lo constituye un extracto de células COS-7 transfectadas con el ADNc 4. **B)** Inmunoprecipitación de un extracto total de células COS-7 con el antisero 1118 analizada mediante *Western Blot* con el anticuerpo anti-4.1R 10b. Se muestra el extracto total (T), la fracción no unida (NU) y la fracción unida (U). **C)** La inmunoprecipitación detallada en el apartado B se analizó mediante *Western Blot* con el antisero 1118. **D)** Inmunoprecipitación de un extracto de células COS-7 con el anticuerpo 10b y análisis mediante *Western Blot* con el suero inmune 1118.

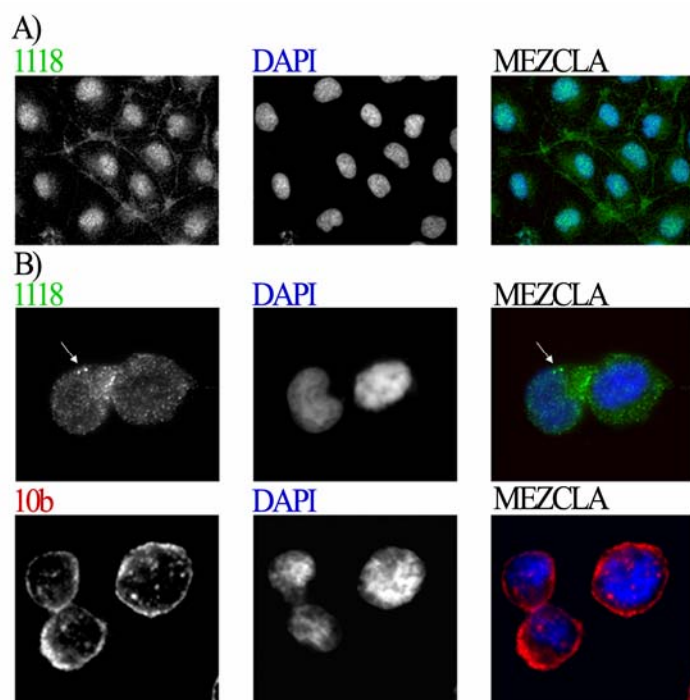
## 2.9. Localización de las isoformas de 4.1R sin CTD en diferentes líneas celulares

### 2.9.1. Localización en células interfásicas.

Se analizó la localización de las isoformas 4.1R sin CTD en diferentes líneas celulares mediante inmunofluorescencia con el anticuerpo 1118. En células interfásicas apenas se detecta proteína salvo en algunas localizaciones concretas que se describen a continuación:



*Localización en la membrana plasmática, colocalizando con la actina cortical.* En las líneas celulares COS-7, ECV304 y MOLT-4 las isoformas de 4.1R sin CTD se localizan en la zona de membrana plasmática mediante la cual contactan dos células (fig. 41A). El análisis de células MOLT-4 mediante inmunofluorescencia con el anticuerpo anti-4.1R 10b, que detecta a todas las isoformas de 4.1R (con o sin dominio CTD), producía, sin embargo, la tinción de toda la membrana plasmática (comparar fig. 41B panel superior e inferior).



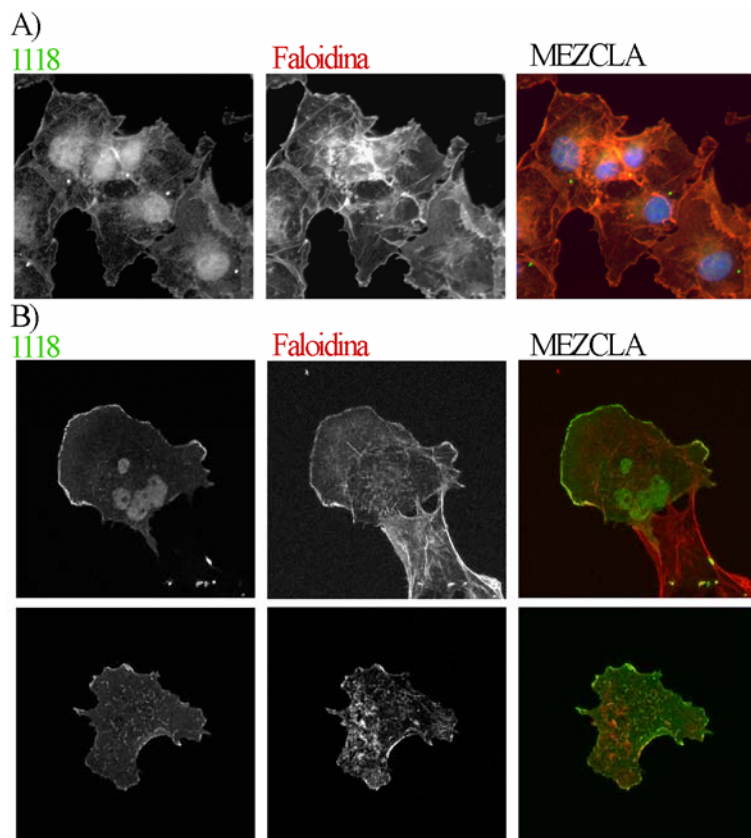
**Figura 41. Localización de las isoformas de 4.1R sin dominio CTD en membrana plasmática.**

**A)** Inmunofluorescencia de células ECV304 con los anticuerpos 1118 y DAPI. Nótese la tinción en membrana plasmática.

**B)** Paneles superiores: inmunofluorescencia con el anticuerpo 1118 de células MOLT-4.

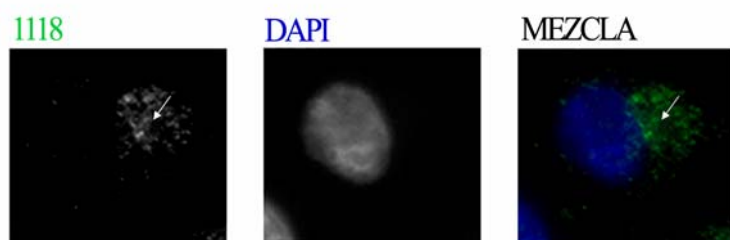
Paneles inferiores: inmunofluorescencia de células MOLT-4 con el anticuerpo 10b.

Ya que las isoformas de 4.1R sin el dominio CTD mantienen el dominio de unión a actina descrito para el resto de isoformas de 4.1R, se analizó una posible colocalización con la actina. Como se muestra en la figura 42A las isoformas de 4.1R sin CTD endógenas colocalizan con la actina cortical. Esta colocalización se observa también cuando los ADNc para las isoformas aisladas en este trabajo se expresan en células COS-7. En este caso, se observaba una colocalización de ambas proteínas en los lamelipodios y en las ondulaciones de la membrana (fig. 42B).



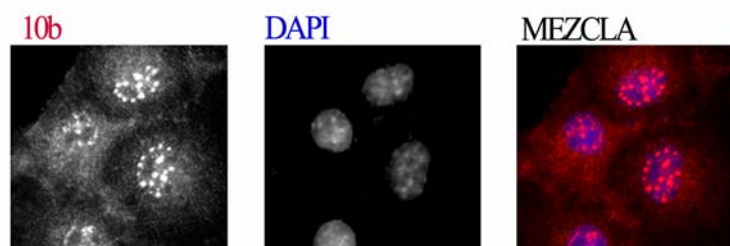
**Figura 42. Colocalización de las isoformas 4.1R sin dominio CTD con la actina cortical.** A) Inmunofluorescencia con el anticuerpo 1118 y con faloidina de células COS-7. B) Inmunofluorescencia con el anticuerpo 1118 y con faloidina de células COS-7 transfectadas con los ADNc 29 (paneles superiores) y 4 (paneles inferiores).

*Localización de dos puntos cercanos al núcleo en células fijadas con metanol.* En muchas células COS-7 y en MOLT-4 se observaba la tinción de dos puntos cercanos al núcleo. En el caso de células COS-7 se observaba además un patrón muy tenue filamentoso alrededor de estos dos puntos (flechas fig. 43).



**Figura 43. Localización de las isoformas de 4.1R sin CTD en dos puntos cercanos al núcleo.** Inmunofluorescencia con el anticuerpo 1118 de células COS-7. La flecha marca la localización de los dos puntos.

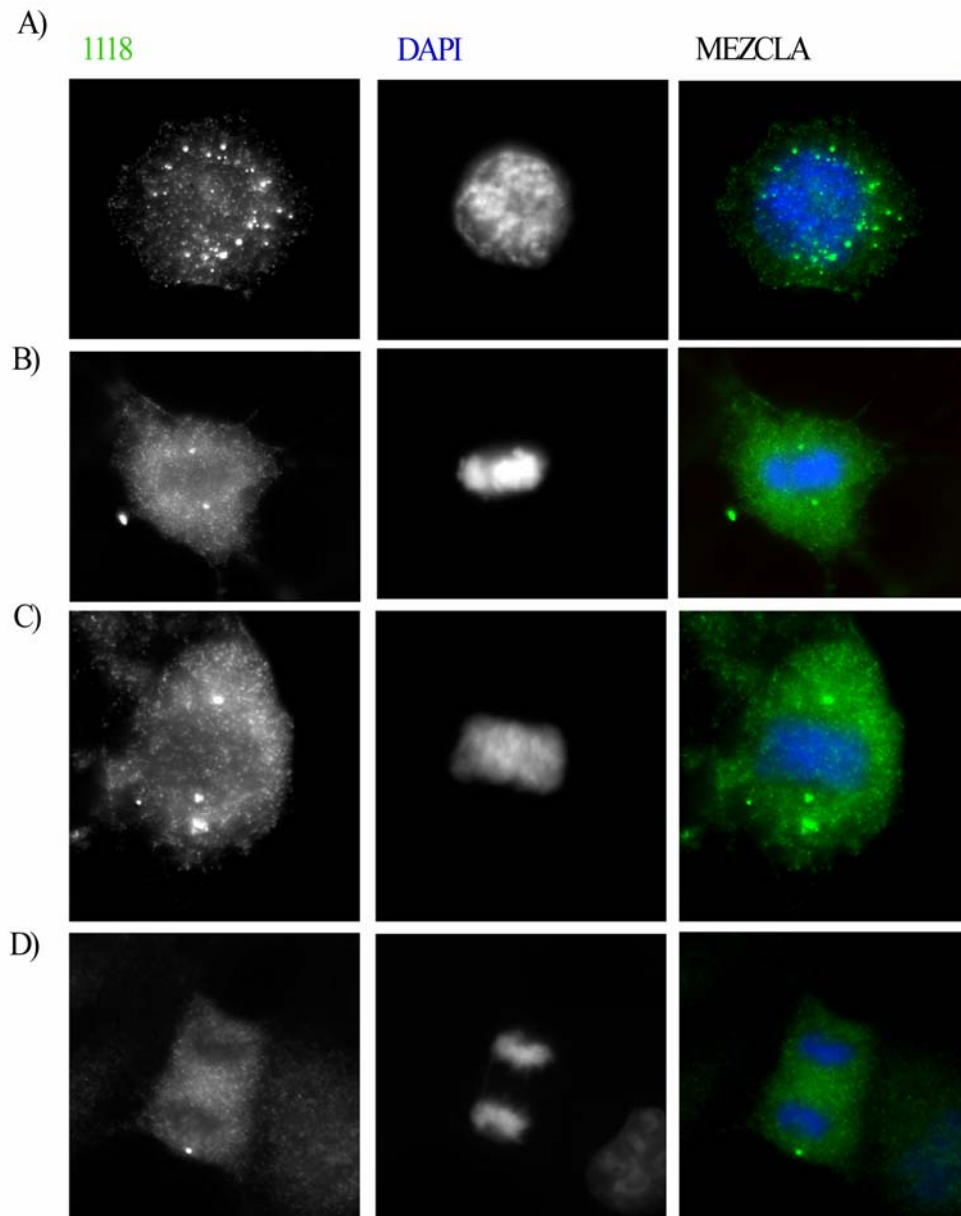
En ninguna de las inmunofluorescencias de las líneas celulares analizadas (MDCK, HEPG2, COS-7, HeLa, ECV304, MOLT-4 y Jurkat) se detectó a las isoformas de 4.1R sin dominio CTD en la matriz nuclear, donde se localizan las isoformas de 4.1R con dominio CTD (fig. 44).



**Figura 44. Localización de las isoformas de 4.1R con dominio CTD en la matriz nuclear.** Inmunofluorescencia de células COS-7 con el anticuerpo 10b.

### 2.9.2. Localización en células en división.

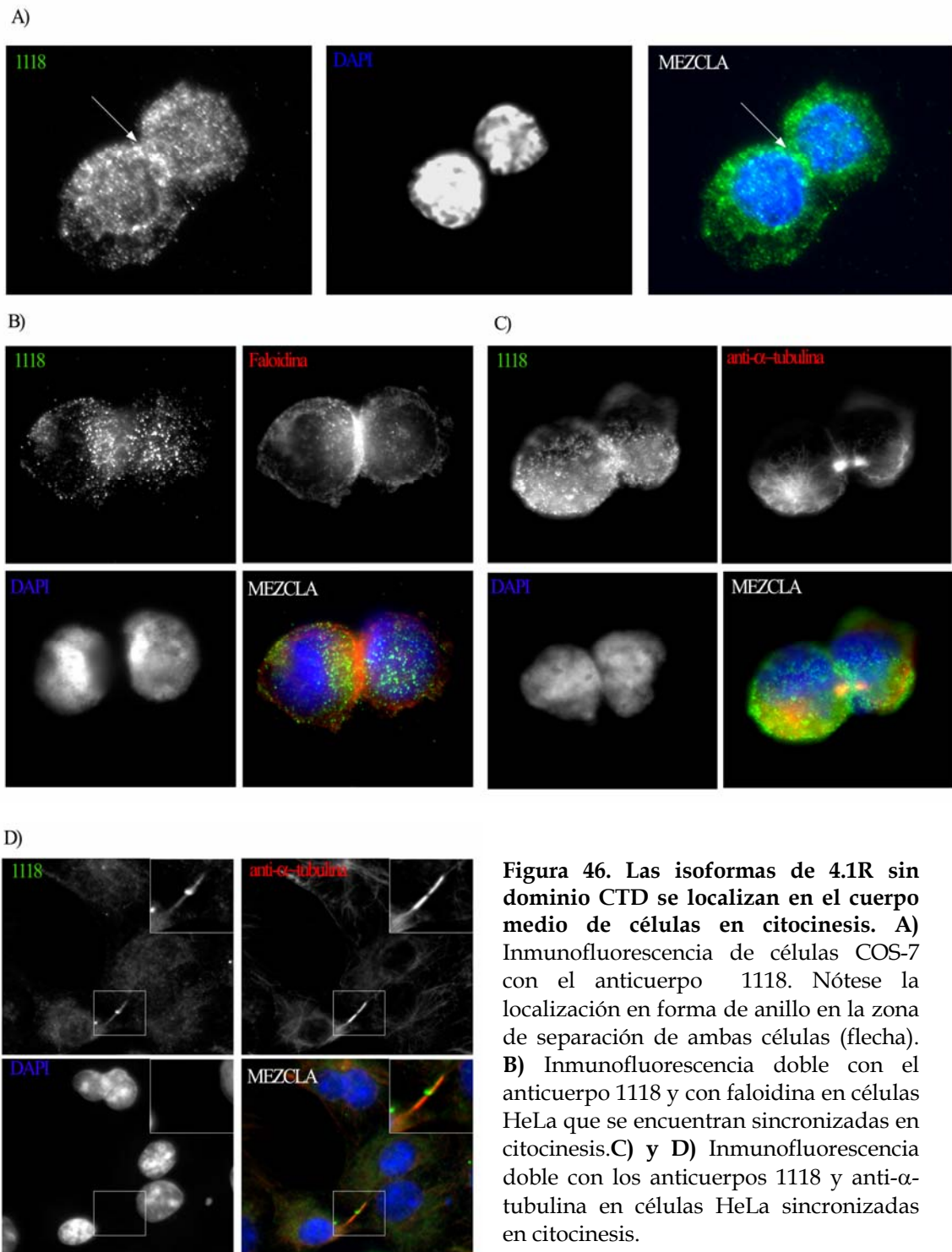
Se estudió la distribución de las isoformas de 4.1R sin dominio CTD durante la división celular mediante inmunofluorescencia con el anticuerpo 1118 en células COS-7 y HeLa. En la figura 45 se muestran imágenes representativas de lo observado. La tinción que presentan estas células varía en las distintas etapas de la división. El patrón obtenido es similar al descrito para las isoformas de 4.1R con dominio CTD (Krauss et al., 1997), con la diferencia de que las isoformas de 4.1R sin CTD no se localizan en el huso mitótico.



**Figura 45. Localización de las isoformas de 4.1R sin CTD a lo largo del ciclo celular.** Se muestran fotos representativas de inmunofluorescencias con el anticuerpo 1118 y DAPI en distintas etapas del ciclo celular. **A)** Célula HeLa en profase. **B)** Célula COS-7 en metafase, nótese los centrosomas **C)** Célula COS-7 en anafase temprana. **D)** Célula COS-7 en anafase tardía.

Al comenzar la citocinesis se observa como las isoformas de 4.1R sin dominio CTD se localizan en la zona de separación de ambas células, formando un anillo (fig. 46A). Antes de que las isoformas de 4.1R sin dominio CTD se sitúen en dicha zona, se acumulan dejando un patrón muy punteado, en la zona cercana al anillo contráctil de actina (fig. 46B).

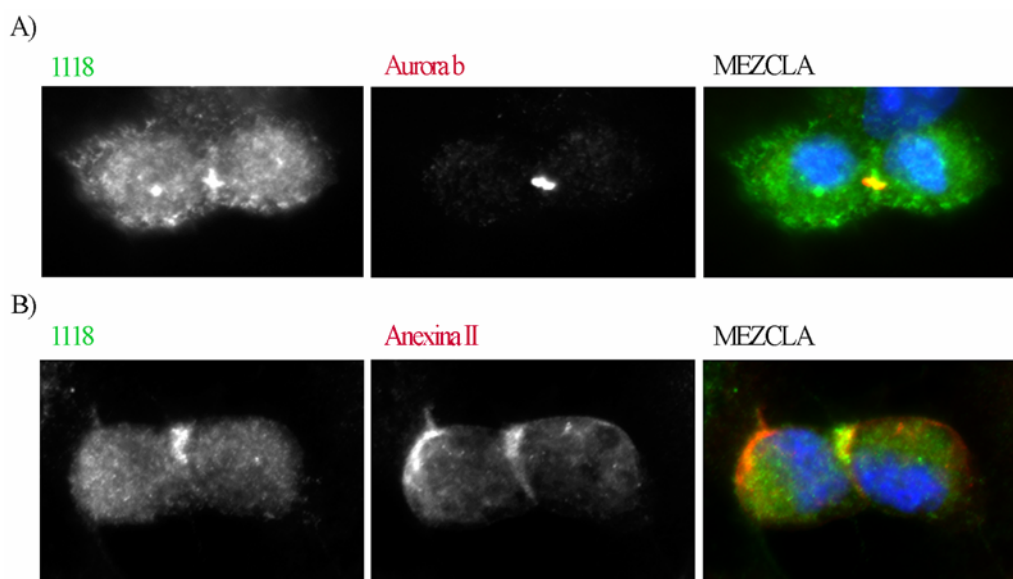
A medida que las células se separan las isoformas de 4.1R sin dominio CTD se localizan en el espacio intermedio que deja la  $\alpha$ -tubulina del cuerpo medio (fig. 46C y D).



**Figura 46.** Las isoformas de 4.1R sin dominio CTD se localizan en el cuerpo medio de células en citocinesis. A) Inmunofluorescencia de células COS-7 con el anticuerpo 1118. Nótese la localización en forma de anillo en la zona de separación de ambas células (flecha). B) Inmunofluorescencia doble con el anticuerpo 1118 y con faloidina en células HeLa que se encuentran sincronizadas en citocinesis. C) y D) Inmunofluorescencia doble con los anticuerpos 1118 y anti- $\alpha$ -tubulina en células HeLa sincronizadas en citocinesis.

## 2.10. Las isoformas de 4.1R sin CTD colocalizan con aurora B y con anexina II en células en citocinesis.

Los resultados anteriores indican que las isoformas de 4.1R sin dominio carboxilo terminal presentan un gran dinamismo a lo largo de la división celular. Se analizó si existía alguna colocalización con proteínas que regulan la división celular como son aurora B y Bub-1. El estudio de inmunofluorescencias dobles permitió establecer que Bub-1 no colocaliza con los puntos en los que se localiza la isoforma 4.1R sin dominio CTD en ninguna fase de la mitosis (no se muestra). Sin embargo, el estudio de células HeLa sincronizadas en citocinesis permitió establecer que las isoformas de 4.1R sin dominio CTD colocalizan con aurora B en esta etapa del ciclo celular (fig. 47A). Se estudió también la distribución entre 4.1R y anexina II, proteína relacionada con la formación del cuerpo medio que separa dos células que se encuentran en citocinesis. Esta proteína se sitúa en el mismo punto que marcan las isoformas 4.1R sin dominio CTD (fig. 47B). Así, todos estos resultados sugieren que las isoformas de 4.1R sin dominio CTD son un componente importante del cuerpo medio de células en citocinesis.



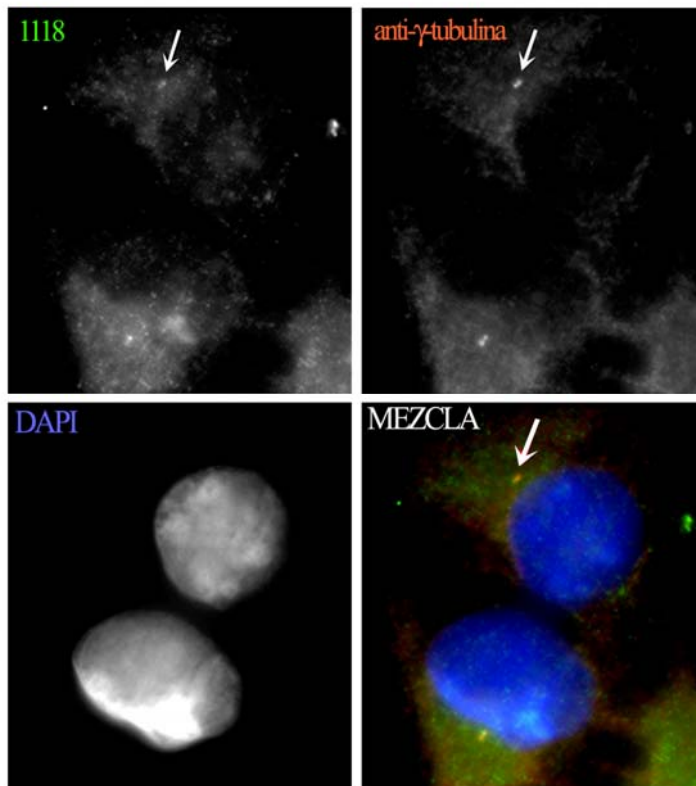
**Figura 47. Las isoformas de 4.1R sin CTD colocalizan durante la citocinesis con aurora B y anexina II. A)** Inmunofluorescencia doble con los anticuerpos 1118 y anti-aurora B en células HeLa sincronizadas en citocinesis. **B)** Inmunofluorescencia doble con los anticuerpos 1118 y anti-anexina II en células HeLa sincronizadas en citocinesis

## 2.11. Las isoformas de 4.1R sin dominio CTD se localizan en el centrosoma de células COS-7 y MOLT-4.

Se analizaron mediante microscopía confocal inmunofluorescencia dobles con los anticuerpos 1118 y anti- $\gamma$ -tubulina de células COS-7 (fig. 48). Se determinó así que los dos puntos en los que se localizaban las isoformas de



4.1R sin dominio CTD en células interfásicas colocalizaban con el centrosoma.

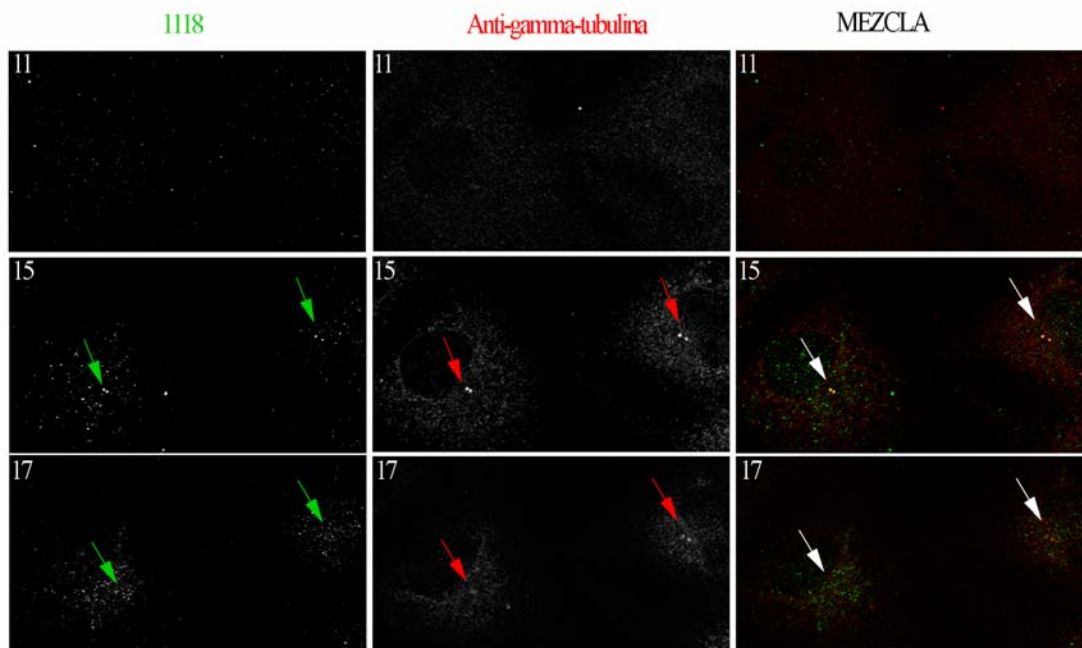


**Figura 48. Colocalización de las isoformas de 4.1R CTD y el centrosoma.**

Inmunofluorescencia doble con el anticuerpo 1118 y anti- $\gamma$ -tubulina en células COS-7. Las flechas señalan el centrosoma. Nótese una colocalización de las isoformas de 4.1R sin dominio CTD con el centrosoma.

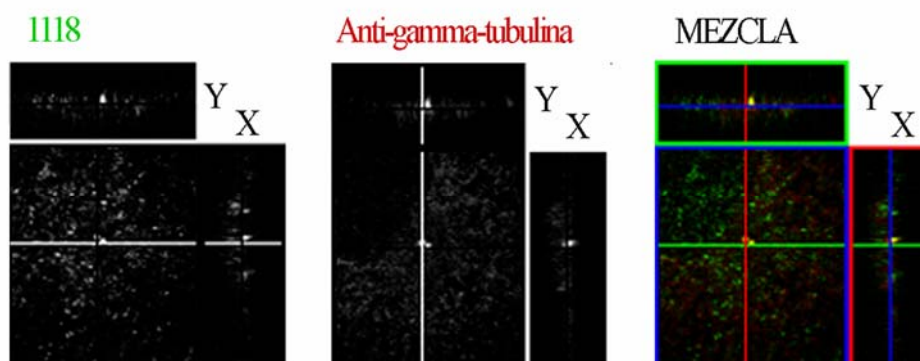
El estudio de diferentes planos indicó que esta colocalización no ocurre a lo largo de toda la estructura del centrosoma. En la figura 49 se muestran tres planos de los 28 tomados en una de las inmunofluorescencias dobles de células COS-7. Se puede observar que existe una colocalización en el plano 15 (señalados con una flecha) pero no en el plano 17 en el que se detectaba el centrosoma pero no a las isoformas de 4.1R sin dominio CTD.

Se eligió uno de estos planos en los que el centrosoma y las isoformas de 4.1R sin dominio CTD colocalizan para realizar cortes ortogonales. Estos cortes permitieron establecer que a lo largo de los dos centriolos se observa una colocalización con las isoformas de 4.1R sin dominio CTD (fig. 50). Por lo tanto, los distintos cortes en el eje X e Y a lo largo de este plano permiten concluir que existe colocalización de las isoformas 4.1R sin dominio carboxilo en determinadas zonas de los centriolos.



**Figura 49. Las isoformas de 4.1R sin dominio CTD colocalizan con el centrosoma en algunos planos.** Inmunofluorescencia doble de células COS-7 con los anticuerpos 1118 y anti- $\gamma$ -tubulina analizadas mediante microscopía confocal. Se tomaron 28 planos de los que se muestran tres, el plano 11, 15 y 17 (indicados en blanco en cada panel). Los planos de la izquierda muestran la tinción para el anticuerpo 1118, los centrales para el anticuerpo anti- $\gamma$ -tubulina y los de la derecha la mezcla. Las flechas muestran la localización de los centrosomas en el plano 15 y 17.

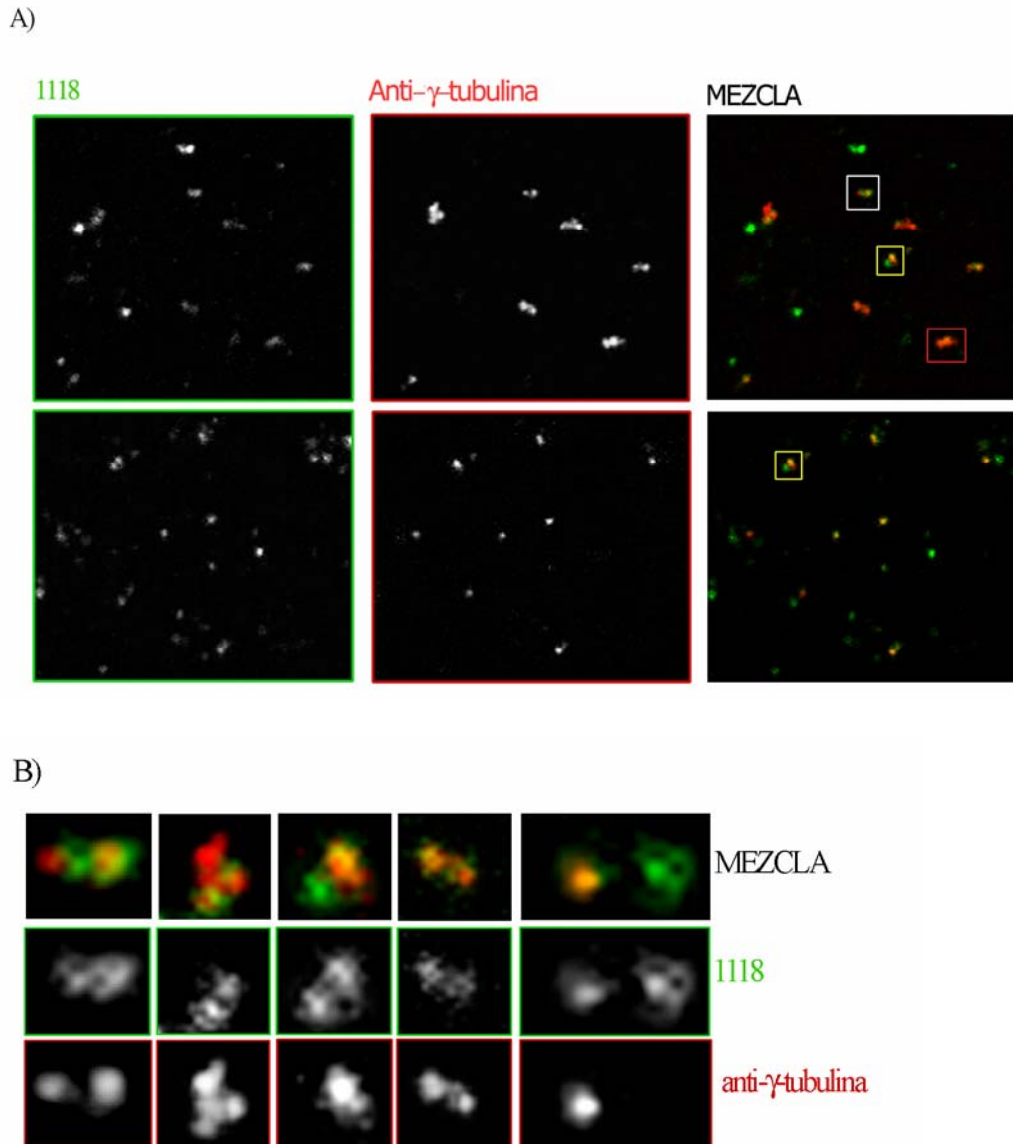
Se aislaron centrosomas de células MOLT-4 y se analizaron mediante inmunofluorescencias dobles con los anticuerpos anti- $\gamma$ -tubulina y 1118 por microscopía confocal. En los diferentes planos obtenidos la mayoría de los centrosomas mantenían asociadas a las isoformas de 4.1R sin dominio CTD.



**Figura 50. Colocalización de las isoformas sin el dominio CTD con el centrosoma.** Se realizaron inmunofluorescencias dobles de células COS-7 con el anticuerpo 1118 y el anticuerpo anti- $\gamma$ -tubulina y se analizaron mediante microscopía confocal. Se eligió uno de los planos en los que colocalizaban estas proteínas para realizar cortes ortogonales. Se indica el corte en el eje Y e X. Nótese la colocalización entre  $\gamma$ -tubulina y las isoformas 4.1R sin dominio CTD.



En algunos casos, estas isoformas presentaban una colocalización exacta con el centrosoma mientras que en otros no colocalizaban totalmente (cuadro rojo y blanco de la figura 51A respectivamente). También se detectaron centrosomas en los que el anticuerpo anti- $\gamma$ -tubulina detectaba un único punto y el anticuerpo 1118 dos (fig. 51A cuadro amarillo).

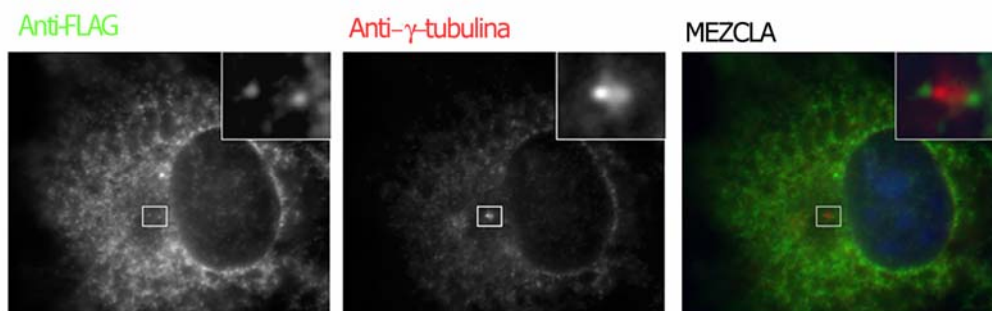


**Figura 51. Análisis de las isoformas de 4.1R sin dominio CTD en centrosomas aislados de células MOLT-4. A)** Inmunofluorescencia doble con los anticuerpos anti- $\gamma$ -tubulina (rojo) y 1118 (verde) de centrosomas aislados de células MOLT-4. Se han señalado con un cuadro rojo un ejemplo en el que ambas proteínas colocalizan, con un cuadro amarillo un ejemplo en el que se observan dos puntos con el anticuerpo 1118 mientras que la  $\gamma$ -tubulina se distribuye en un sólo punto y con un cuadro blanco un ejemplo en el que la proteína 4.1R sin CTD localiza solo parcialmente con la tubulina. **B)** Se han ampliado algunos de los centrosomas del apartado A. Nótese como la tinción con el anticuerpo 1118 parece mostrar una disposición concreta en la que se rodea a los centriolos (última columna de paneles).

Mediante el uso de programas para el análisis de imágenes se ampliaron fotos de diferentes planos focales de los centrosomas (fig. 51B). Esto permite sugerir que en muchas ocasiones las isoformas de 4.1R sin dominio CTD están localizadas alrededor de los centriolos, formando dos estructuras circulares. Este tipo de estructura se observa en diferentes centrosomas y en distintas orientaciones (segunda y tercera columna fig. 51B).

Se analizó mediante *Western Blot* con el anticuerpo 1118 la fracción de centrosomas aislados de células MOLT-4. Se observó que en estos centrosomas existe una isoforma de 4.1R con una movilidad electroforética de 110 kDa (no mostrado).

Se estudió si alguna de las isoformas de 4.1R sin dominio CTD aisladas en este trabajo colocalizaba con el centrosoma, cuando se expresaban en células COS-7 a partir de los ADNc correspondientes. Estas células se analizaron mediante inmunofluorescencias dobles con los anticuerpos anti  $\gamma$ -tubulina y anti-FLAG. En aquellas células transfectadas con los ADNc para las isoformas 8 y 32 se observa que estas isoformas presentan una localización similar a la observada para las isoformas endógenas en los centrosomas de MOLT-4, localizándose alrededor de este orgánulo (fig. 52).



**Figura 52. Localización de las isoformas de 4.1R sin CTD expresadas exógenamente en células COS-7 respecto al centrosoma.** Inmunofluorescencias dobles con los anticuerpos anti-FLAG (verde) y anti- $\gamma$ -tubulina (rojo) en células COS-7 transfectadas con el ADNc 32 marcado con el epítipo FLAG en el extremo 5'.

## Discusión

## DISCUSIÓN

La clonación de ADNc para 4.1R ha permitido establecer la gran capacidad que presenta el gen *EPB41* para generar diferentes isoformas de esta proteína. Pronto se estableció que esta variabilidad se debía al procesamiento alternativo al que era sometido su pre-ARNm y que éste dependía del tejido y del momento del desarrollo. Nuestro laboratorio ha contribuido a establecer los diferentes patrones de procesamiento alternativo que sufre el pre-ARNm de esta proteína, mediante el aislamiento de muchos de sus ADNc.

A pesar de que la gran variabilidad de isoformas de 4.1R se ha atribuido al procesamiento alternativo de su pre-ARNm, son pocos los estudios que han establecido la influencia de otros mecanismos en la generación de isoformas de esta proteína. En este trabajo se plantea la implicación de otros mecanismos que pueden influir en la generación de isoformas de 4.1R.

### **1. UN MECANISMO DE TRADUCCIÓN INTERNA PERMITE LA GENERACIÓN DE ISOFORMAS DE PESO MOLECULAR INTERMEDIO DE 4.1R.**

#### **1.1. Dos isoformas de 4.1R son expresadas desde un único ADNc.**

La expresión de los genes eucariotas suele ocurrir a partir de ARNm monocistrónicos, por lo que sólo generan una proteína. Sin embargo, existen en la actualidad ejemplos de genes eucariotas policistrónicos (Blumenthal, 2004). Uno de los primeros ejemplos descritos lo constituye el ARNm del factor de expresión y diferenciación 1 (GDF-1) cuyo ARNm codifica también para la proteína UOG-1 (Lee, 1991). En ocasiones estos ARNm policistrónicos expresan isoformas de una misma proteína, como ocurre en el caso de la proteína Hairless de *Drosophila melanogaster* (Maier et al., 2002).

En este trabajo, los resultados obtenidos mediante la transcripción y traducción *in vitro* de distintos ADNc para la proteína 4.1R pusieron de manifiesto que algunos generan dos proteínas. Ambas son isoformas de la proteína 4.1R, una de ellas perteneciente al grupo de isoformas de peso molecular alto (135 kDa) y la otra al grupo de peso molecular intermedio (80kDa). El análisis de la composición exónica de los ADNc utilizados indicó que aquellos capaces de generar dos isoformas contenían el exón 2' y el 4. Los mismos resultados se obtuvieron mediante ensayos de expresión *in vivo* de estos ADNc. Los datos disponibles hasta el momento, atribuyen la producción de isoformas de peso molecular intermedio de 4.1R a la existencia de ARNm que carecen del exón 2', y por lo tanto del AUG1. La traducción de estas isoformas ocurre a partir del AUG2 presente en el exón 4. En este trabajo se describe por primera vez que la producción de isoformas de 4.1R de peso molecular intermedio tiene lugar también a partir de los ARNm que contienen el exón 2'.

Diversos mecanismos pueden explicar la producción de diferentes proteínas o isoformas de las mismas, a partir de un único ADNc. Entre ellos, como se detalló en la introducción de este trabajo, se encuentran procesos que actúan a

nivel de la transcripción y del procesamiento del pre-ARN o en el inicio de la traducción de éste.

## **1.2. Una traducción interna está implicada en la generación de la isoforma de 4.1R de peso molecular intermedio.**

Un proceso de proteólisis puede generar dos isoformas de una proteína a partir de la información contenida en un único ADNc. Un ejemplo de esto sucede en el factor de crecimiento del endotelio vascular (VEGF-D) (Baldwin et al., 2001), en el que diferentes isoformas son generadas mediante proteólisis. Los resultados obtenidos en este trabajo, utilizando mutantes que impedían la producción de la isoforma de 4.1R de peso molecular alto, descartan la implicación de un proceso de proteólisis en la generación de dos isoformas de esta proteína a partir de un único ADNc.

Otro mecanismo responsable de la generación de dos isoformas de una misma proteína a partir de un único ADNc es la existencia de promotores alternativos en el pre-ARN. Existen diferentes ejemplos de este mecanismo como es el caso del pre-ARN que codifica para las diferentes isoformas de la tropomiosina (Gunning et al., 2005). Los ADNc para 4.1R utilizados en el desarrollo de este trabajo son expresados a partir del promotor del vector en el que están clonados. Los resultados obtenidos con construcciones en las que se eliminó este promotor excluyen la existencia de más promotores dentro de la secuencia de los ADNc, descartando que sea éste el mecanismo que genera las dos isoformas de 4.1R.

La traducción interna de un ARNm puede explicar la generación de dos isoformas de una proteína. Este mecanismo es muy utilizado en virus y ha sido descrito también para algunos ARNm eucariotas (Bonnal et al., 2003). Aunque hubo mucha controversia acerca de la existencia de este mecanismo en estos últimos (Kozak, 2001; Kozak, 2002; Kozak, 2003), la lista de los ARNm eucariotas en los que se da una traducción interna ha aumentado considerablemente en los últimos años. Los resultados de la expresión de un vector bicistrónico que portaba la secuencia nucleotídica de 4.1R comprendida entre el ATG1 y el ATG2, indican que dicha secuencia es capaz de dirigir una traducción interna. En ocasiones la expresión de los dos cistrones de un vector bicistrónico no se debe a una traducción interna sino a la existencia de ARNm monocistrónicos generados bien por procesamiento, bien por la transcripción de promotores alternativos (Han and Zhang, 2002; Liu et al., 2005). En este trabajo se ha descartado la existencia de promotores crípticos y la de ARNm monocistrónicos en la generación de las dos isoformas de 4.1R desde un único ADNc. Todos estos resultados indican que la producción de dos isoformas de 4.1R tiene lugar a través de una traducción interna.

## **1.3. Traducción interna. *Leaky scanning*, *shunting* o IRES.**

El modelo del rastreo del ribosoma, establecido para el inicio de la traducción de un ARNm, es un proceso complejo que como detallamos en la introducción de este trabajo depende no sólo de multitud de factores sino también de la secuencia en la que se encuentra el triplete iniciador. Se han descrito modelos

alternativos al del rastreo del ribosoma para el inicio de la traducción, entre los que se encuentran el *shunting*, el *leaky scanning* y los IRES o sitios de entrada directa para el ribosoma. De estos mecanismos, la reiniciación no puede originar las dos isoformas de 4.1R, ya que la secuencia de una de estas isoformas se encuentra contenida en la secuencia que codifica para la otra isoforma.

En principio, el *leaky scanning* utilizado en general por aquellos ARNm que contienen codones de iniciación que no se encuentran en un contexto de secuencia óptimos, podría explicar la generación de dos isoformas de 4.1R a partir de un único ADNc. Los tripletes para el inicio de la traducción en el ARNm de 4.1R, ATG1 y ATG2, no están en un contexto favorable, ya que ninguno presenta una guanina en la posición +3. Este proceso ha sido descrito en diferentes ARNm como por ejemplo en el que codifica para (Calkhoven et al., 2000). La producción de dos isoformas de 4.1R desde un único ADNc también podría llevarse a cabo mediante *shunting*. Este mecanismo lo presenta el ARNm que codifica para la  $\beta$ -secretasa, que genera mediante este proceso dos isoformas de esta proteína (Meyer et al., 2003). Por último, una traducción dependiente de IRES, (secuencia interna de entrada del ribosoma) como la que presentan los ARNm que codifica para PITSLRE (Cornelis et al., 2000), también podría explicar la producción de dos isoformas de 4.1R a partir de un único ADNc.

Para que tenga lugar una iniciación mediante *leaky scanning* o *shunting* es necesario que el complejo de preiniciación de la traducción reconozca la estructura CAP del extremo 5' del ARNm y que rastree la región 5' del ARNm anterior al codón de iniciación. Sin embargo, en una traducción dependiente de IRES el ribosoma se une de forma directa a una secuencia interna del ARNm sin necesidad de reconocer la estructura CAP, ni de rastrear la zona anterior al codón de iniciación.

Se introdujo una secuencia altamente estructurada ( $\Delta G^\circ$  de -57 kCal/mol) en la región 5' de la secuencia de 4.1 que portaba el vector bicistrónico. Al llegar a esta secuencia el ribosoma no puede avanzar y el rastreo del ARNm queda impedido, de forma que no se produce la traducción dependiente de rastreo del segundo cistron. A pesar de esto, se obtuvo expresión del segundo cistron del vector. Esto indica que el ribosoma se ancla directamente a la secuencia de 4.1 sin necesidad de rastrear la secuencia 5' anterior. Por lo tanto el mecanismo mediante el cual se lleva a cabo la traducción interna de isoformas de 4.1R de peso molecular intermedio es independiente del rastreo y está mediado por un IRES.

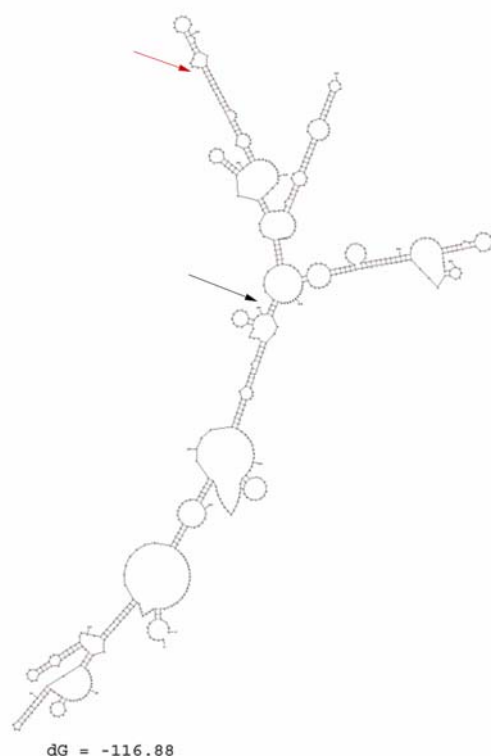
#### **1.4. La traducción interna del ADNc para 4.1R es dependiente de la línea celular y específica de la secuencia del ADNc.**

El estudio de la expresión del vector bicistrónico indicó que no todas las líneas celulares testadas en este trabajo son capaces de generar la expresión del segundo cistron. Así, mientras que en las líneas HeLa (humana) y COS-7 (mono verde africano) se produce la expresión del segundo cistron, en la línea MDCK (perro) no se ha observado una expresión de éste.

La comparación de la secuencia de 4.1R humana clonada en el vector bicistrónico con la correspondiente secuencia de perro muestra un gran número

de diferencias, mientras que la comparación de la secuencia humana con la de mono presenta muchas menos. Este dato indica, bien una especificidad de la secuencia capaz de dirigir la traducción interna, bien la necesidad de algún factor presente tan solo en algunos tipos celulares, o bien ambas. La necesidad de factores para la traducción interna mediada por IRES ha sido previamente descrita (Pilipenko et al., 2000) (Holcik et al., 2003; Pickering et al., 2003), denominándose a estos factores ITAFs (IRES trans acting factors). El hecho de que los ADNc para 4.1R generen dos isoformas de esta proteína en un lisado de reticulocitos hace pensar que no sean necesarios factores adicionales a los presentes en dichos lisados para su traducción. Por lo tanto, parece más probable que un cambio de la secuencia del IRES sea la responsable de que no se observe la expresión del segundo cistron en las células MDCK.

El análisis llevado a cabo con diferentes fragmentos de la secuencia de 4.1R comprendida entre el ATG1 y el ATG2 clonados en los vectores bicistrónicos, muestra que el acortamiento de la región 5' influye negativamente en la expresión del segundo cistron. La mayor producción del segundo cistron se obtuvo para la construcción que portaba la secuencia de 4.1R comprendida entre el ATG1 y el ATG2. La eliminación de 153 nucleótidos del extremo 5' de la secuencia de 4.1R disminuyó la relación del segundo cistron sobre el primero de 1.75 a 1.50. Sin embargo, la eliminación de otros 76 nucleótidos provoca que esta relación disminuya hasta 0.3. Esto parece indicar que dentro de esos 76 nucleótidos podrían encontrarse secuencias importantes para la actividad del IRES de 4.1R.



**Figura 52.** Modelo de la estructura secundaria del fragmento de ARNm de 4.1R comprendido entre el ATG1 y ATG2. La flecha negra indica la zona 5' en la que comienza la secuencia de 4.1R clonada en la construcción DsRed-4.1Δ153-EGFP y la flecha roja la de la construcción DsRed-4.1Δ229-EGFP. Se indica también, la energía libre estimada para este modelo (-116,8 Kcal/mol).



Se analizó el plegamiento de la secuencia comprendida entre el ATG1 y ATG2 del ARN de 4.1R mediante el programa mfold, ya que la estructura secundaria que presentan algunos IRES están relacionadas con la actividad de estos (Baird et al., 2006). Como se puede observar en la figura 52, la eliminación de la zona anterior a la señalada con una flecha negra, responsable de la disminución de la relación del segundo cistrón respecto del primero de 1,75 a 1,5, no elimina muchos de los lazos del ARN. Sin embargo, la eliminación de la secuencia anterior a la señalada con una flecha roja, elimina la zona en la que se localizan la mayor parte de lazos de ARN. Esto podría explicar que en la construcción en la que se eliminan otros 76 nucleótidos de la secuencia de 4.1R, clonada en el vector bicistrónico, se produzca una relación del segundo cistrón respecto del primero de 0.3.

La construcción E4-4.1R que portaba la misma secuencia de 4.1R que la del vector bicistrónico (que reduce la relación del segundo cistrón respecto del primero a 0.3) y que mantiene el resto de secuencia-3' de 4.1R, generó la isoforma de peso molecular intermedio tanto *in vivo* como *in vitro*. Este dato aparentemente contradictorio puede explicarse teniendo en cuenta bien, que la cercanía de los ATG iniciadores (ATG1' y ATG2) en la construcción E4-4.1R puede favorecer el proceso, o bien a que sea la secuencia posterior al ATG2 de 4.1R de dicha construcción la que lo favorezca. Otro factor que puede influir en este resultado es la presencia del cistrón que codifica para DsRed en el ARN mensajero generado a partir del vector bicistrónico. Este ARN no porta únicamente el IRES de 4.1R sino que presenta en la zona 5' el cistrón que codifica para DsRed, que podría afectar bien a la estructura, bien a la función de este IRES.

### 1.5. Posibles implicaciones biológicas de un IRES en la secuencia del ADNc para 4.1R.

La expresión de los IRES eucariotas ha sido descrita en aquellos casos en los que la traducción mediante el modelo tradicional por rastreo del ribosoma está inhibida. Esto ocurre en distintas situaciones como son la apoptosis, (Henis-Korenblit et al., 2000; Holcik et al., 1999; Stoneley et al., 2000) la división celular (Pyronnet et al., 2000) (Pyronnet and Sonenberg, 2001) o condiciones de estrés celular provocadas por falta de nutrientes u oxígeno (Stein et al., 1998), por choque osmótico (Allemand et al., 2005) o confluencia de las células (Morrish and Rumsby, 2002) entre otras. Los IRES descritos en estas situaciones permiten la expresión de proteínas necesarias en los procesos citados.

Las células transfectadas bien con el vector bicistrónico, bien con ADNc para 4.1R se sometieron a diferentes condiciones de estrés (choque osmótico, térmico, retirada de suero, tratamiento con estaurosporina y confluencia) sin que se produjese un aumento significativo en la expresión del segundo cistrón o de la isoforma de peso molecular intermedio de 4.1R, respectivamente. La sincronización mediante nocodazol de células transfectadas con el vector bicistrónico tampoco produjo un aumento significativo de la expresión del segundo cistrón con respecto a las células control. Teniendo en cuenta que la obtención de la isoforma de 4.1R de peso molecular intermedio desde un único

ADNc variaba en los diferentes experimentos, pero que siempre se producía, es poco probable que sea una condición de estrés la que regule la expresión de esta isoforma.

Nuestro grupo ha descrito recientemente la capacidad de las isoformas de 4.1R de peso molecular intermedio de formar oligómeros (Perez-Ferreiro et al., 2006). A pesar de que la región mediante la cual se forman dichos oligómeros está presente en los tres tipos de isoformas de 4.1R (peso molecular alto, intermedio y bajo) las isoformas de peso molecular alto e intermedio no se asocian debido a que la conformación que adoptan oculta la zona necesaria para la interacción. Se planteó la posibilidad de que las dos isoformas de 4.1R de peso molecular alto e intermedio, al estar traducidas desde un mismo ADNc, pudieran interaccionar formando oligómeros antes de que adoptaran la conformación que oculta la zona de interacción. Esta posible interacción no ha sido detectada ni mediante inmunoprecipitaciones ni mediante el uso de gradientes de sacarosa. Existen ejemplos como el de la proteína THI1 (Chabregas et al., 2003) en el que una de las dos isoformas de una proteína, generada por traducción interna, presenta una localización diferente. Por lo tanto, una posibilidad para la expresión de las dos isoformas de 4.1R desde un único ADNc sería su implicación en algún proceso que se llevara a cabo en distintas zonas de la célula.

En cualquier caso parece que las isoformas de 4.1R de peso molecular intermedio deben ser muy importantes para la célula ya que la producción de éstas se puede dar mediante la traducción interna de los ARNm para la isoformas de peso molecular alto, pero también desde sus propios ARNm. No está establecido cual es el mecanismo que permitiría la traducción de aquellos ARNm monocistrónicos que codifican para las isoformas de 4.1R de peso molecular intermedio. Hay que señalar que hasta la fecha, todos estos ARNm monocistrónicos mantienen el exón 2 (475 pb) a pesar de que éste no se traduce. Dada la longitud de esta secuencia, parece improbable un inicio de la traducción de estos ARNm mediante el modelo general por rastreo del ribosoma. Hay que hacer notar que la región que hemos identificado como el IRES de 4.1R se encuentra también en los ARNm que codifican para las isoformas de peso molecular intermedio. Por lo tanto, es posible que la traducción de estos ARNm esté ocurriendo también mediante un IRES.

#### **1.6. El IRES de 4.1R se encuentra en la región codificante.**

Son muchos los IRES que se han descrito en los últimos años, localizados en ARNm codificantes para proteínas muy diversas (Vagner et al., 2001). A pesar de esto, son escasos los ejemplos de ARNm en los que el IRES se localiza en la zona codificante para otra proteína. El primer IRES dentro de la secuencia codificante de un ARNm eucariota fue descrito en el año 2000 en el ARNm codificante para la proteína PITSLRE también denominada CDK11, (quinasa 11 dependiente de ciclina) (Cornelis et al., 2000). Este ARNm genera dos isoformas, la menor de ellas mediante la existencia de un IRES localizado en la secuencia codificante de la isoforma de mayor peso molecular. La isoforma de menor peso molecular se expresa durante la fase G2/M del ciclo celular. Otros dos ejemplos de IRES contenidos en la región codificante se han descrito para los ARNm de

Notch y el de Hairless (Maier et al., 2002). Esta última es una proteína antagonista de Notch, proteína implicada en la diferenciación y en la muerte celular en metazoos superiores. La expresión de estos IRES también es dependiente de ciclo celular. La expresión de dos isoformas de 4.1R desde ADNc para isoformas de peso molecular alto, descrita en este trabajo, indican que el IRES que permite este proceso se incluye en la pequeña lista de IRES localizados en regiones codificantes.

Como se ha señalado anteriormente, este IRES también está presente en la secuencia de los ADNc para las isoformas de peso molecular intermedio, en donde se comportaría como un IRES no codificante. Es el procesamiento alternativo (*splicing*) del pre-ARNm de 4.1R el que determina si se generan ARNm para isoformas de peso molecular alto o intermedio, mediante la inclusión o exclusión del exón 2', respectivamente. Indirectamente, esta elección determina la utilización de un IRES localizado en la región codificante, en aquellos ARNm de isoformas de peso molecular alto, o de un IRES localizado en una región no codificante en el caso de los ARNm para isoformas de peso molecular intermedio. A pesar de que este tipo de regulación no se ha descrito con anterioridad, existen ejemplos en los que el procesamiento alternativo del preARN modifica la actividad de un IRES, como en el ARNm que codifica para la glutamina sintetasa (Shin et al., 2003).

Los resultados de este trabajo indican que en la expresión de las isoformas de 4.1R está implicado no solo el procesamiento del pre-ARNm, sino también la traducción. Es probable que la regulación de ambos procesos esté interrelacionada permitiendo un patrón de expresión muy complejo de las isoformas de 4.1R.

## 2. ISOFORMAS DE 4.1R SIN DOMINIO CTD.

### 2.1. Nuevo patrón de procesamiento en el pre-ARNm de 4.1R.

El NIH (*The National Institute of Health*, Instituto Nacional de la Salud) ha desarrollado el programa MGC (*Mammalian Gene Collection*, Colección de Genes de Mamífero) con el que se pretende reunir un gran número de ADNc completos. El estudio de estos ADNc permite establecer información acerca de la estructura de los genes y la forma en la que éstos se expresan. Hasta ahora, a pesar de la gran cantidad de isoformas de 4.1R descritas todas ellas comparten el dominio carboxilo CTD diferenciándose en la extensión de su zona amino terminal. Sin embargo, uno de los ADNc (BC039079) depositados en el MGC indica la existencia de una isoforma de 4.1R que carece de este dominio CTD (Strausberg et al., 2002). Los experimentos llevados a cabo en el desarrollo de este trabajo han permitido aislar nuevos ADNc que corroboran la existencia de este tipo de isoformas. Así, todos los ADNc aislados se caracterizan por codificar para isoformas de 4.1R que carecen del dominio CTD. Esto es debido a una señal de poliadenilación presente en la secuencia intrónica comprendida entre los exones 17 y 18 introducida, de forma diferencial, en estos ADNc. Anterior a esta señal de poliadenilación se añaden otros 27 nucleótidos, entre los

que se incluye un codón de parada de traducción, que son eliminados probablemente mediante procesamiento alternativo en las isoformas de 4.1R con dominio CTD.

Se ha determinado que aproximadamente la mitad de los genes humanos poseen sitios de poliadenilación alternativa. Esto puede generar diferentes transcritos y por lo tanto aumentar el número de isoformas que posee una proteína. Las señales de poliadenilación alternativas pueden estar localizadas en *tandem* en un mismo exón, o bien en diferentes exones. Estas últimas dependen del patrón de procesamiento alternativo que siga el ARNm. Dentro de estos mecanismos se encuentran los exones compuestos y los exones eliminados. Los exones compuestos se caracterizan por presentar un sitio 5' de procesamiento alternativo que en ocasiones está silenciado. Cuando esto sucede el exón se comporta como exón 3' terminal, mientras que si es activo da lugar a un exón interno. El otro mecanismo, denominado mecanismo de los exones eliminados, es aquel en el que el exón que contiene uno de los sitios de poliadenilación es eliminado por completo (Edwards-Gilbert et al., 1997).

El ARNm codificante para las isoformas de 4.1R sin CTD presenta un exón compuesto de forma que el sitio 5' de procesamiento alternativo, comúnmente utilizado en el resto de isoformas de 4.1R, no se ha utilizado en este ARNm. Esto trae como consecuencia la incorporación de una señal de poliadenilación que no se encuentra en el resto de ARNm de 4.1R.

## **2.2. Composición exónica de los ADNc para isoformas de 4.1R sin dominio CTD.**

Los once ADNc para isoformas sin CTD aislados en este trabajo provienen de cerebro humano y se han clasificado en cuatro grupos en función de su composición exónica. Uno de ellos, presenta una composición exónica idéntica a la del clon BC039079 mientras que los otros, difieren en cuanto a la presencia o ausencia de los exones 14, 15 y 16. El procesamiento alternativo de estos tres exones ha sido ya descrito para isoformas de 4.1R aisladas de tejido cerebral de rata (Kontogianni-Konstantopoulos et al., 2001b). No se han detectado en este trabajo, ni se tiene conocimiento, de la existencia de isoformas sin dominio CTD que carezcan del exón 2'.

Dos de los clones aislados muestran una delección de tres nucleótidos en el comienzo del exón 13. Este tipo de procesamiento ha sido descrito con anterioridad y se debe a la existencia de sitios aceptores de procesamiento crípticos en este exón que no provocan un cambio de lectura del ARNm aunque eliminan un aminoácido en la proteína que estos generan (Venezia et al., 1998).

## **2.3. Localización de las isoformas de 4.1R sin CTD.**

Mediante el análisis de la composición de las isoformas sin CTD se puede predecir alguna de las interacciones que estas isoformas pueden o no mantener. Las interacciones con las proteínas FKBP13 (Walensky et al., 1998), NuMa, ZO-1, ZO-2, y eIF3-p44 no serían posibles, en principio, ya que los sitios descritos como necesarios para dichas interacciones se localizan en el dominio CTD. Sin

embargo, todas las isoformas mantienen el sitio de interacción descrito para la unión con la actina (Schischmanoff et al., 1995) y con la tubulina (Perez-Ferreiro et al., 2001). Cabe esperar un comportamiento diferencial de aquellas isoformas que pierden el exón 16, en cuanto a la formación del complejo 4.1R-actina-espectrina, ya que en ella intervienen aminoácidos codificados por los exones 16 y 17 (Burns et al., 1983); (Tyler et al., 1980).

La generación de un anticuerpo específico ha permitido estudiar de forma endógena las isoformas de 4.1R sin dominio CTD. Éstas están presentes en muy baja concentración en las líneas celulares estudiadas, siendo necesaria su inmunoprecipitación para detectarlas. Mediante inmunofluorescencia se confirma la baja concentración de estas isoformas, que aparecen concentradas en zonas concretas de la membrana plasmática, en el centrosoma y, en células en citocinesis, en el cuerpo medio de separación de las células hijas. Las isoformas de 4.1R sin dominio CTD, a diferencia de las isoformas que mantienen este dominio, no se localizan en los dominios nucleares de ninguna de las líneas celulares estudiadas (De Carcer et al., 1995).

La presencia de 4.1R en el centrosoma de células de mamífero fue descrita en 1997 mediante microscopía electrónica (Krauss et al., 1997). Nuestro grupo describió posteriormente la existencia de isoformas de 4.1R en centrosomas aislados de linfocitos (Perez-Ferreiro et al., 2004). Sin embargo ningún trabajo ha identificado todavía las isoformas específicas de 4.1R, localizadas en el centrosoma. En este trabajo hemos determinado que las isoformas de 4.1R sin dominio CTD colocalizan con el centrosoma de células COS-7. El aislamiento de centrosomas de linfocitos nos ha permitido determinar también su presencia en dichas preparaciones. La distribución de las isoformas sin dominio CTD en alrededor de los centríolos, es similar a la descrita por Krauss para las isoformas de 4.1R. Cabe señalar que hasta ahora, los estudios realizados para localizar a 4.1R en el centrosoma se han llevado a cabo con anticuerpos contra epítomos mantenidos en las isoformas de 4.1R sin CTD. Por lo tanto, cabe la posibilidad de que las isoformas detectadas hasta ahora en el centrosoma pertenezcan al grupo de 4.1R sin dominio CTD. La presencia de algunas isoformas de 4.1R sin dominio CTD expresadas de forma exógena en el centrosoma apoya esta hipótesis. Si bien, es posible también que ambos tipos de isoformas de 4.1R, con y sin dominio CTD, estén presentes en el centrosoma.

Las isoformas de 4.1R sin dominio CTD presentan una redistribución a lo largo del ciclo celular muy similar a la que presentan las isoformas de 4.1R que mantienen el dominio CTD (Krauss et al., 1997). La diferencia más destacable es que en los experimentos desarrollados en este trabajo las isoformas sin dominio CTD no se han localizado nunca en el huso mitótico.

Datos obtenidos por Krauss en 1997 indican que las isoformas de 4.1R con dominio CTD se localizan en el cuerpo medio de las células en citocinesis. En este trabajo se muestra que las isoformas sin dominio CTD también se localizan en el cuerpo medio, en las que además colocalizan con anexina II y con aurora B. Esto podría sugerir que las isoformas de 4.1R son un componente importante del cuerpo medio de células en citocinesis.

### 3. LOS MECANISMOS QUE GENERAN ISOFORMAS DE 4.1R ACTÚAN A DISTINTOS NIVELES.

Hasta ahora se conocía que la existencia de un gran número de isoformas de 4.1R se debía principalmente al procesamiento alternativo de su pre-ARN. En concreto la elección de uno de sus tres primeros exones excluyentes determina la presencia o ausencia del exón 2', que a su vez determina la producción de una isoforma de 4.1R peso molecular alto o intermedio, respectivamente.

En este trabajo se demuestra, por primera vez, la implicación de un mecanismo de traducción interna en la generación de isoformas de 4.1R de peso molecular intermedio. La posible generación de las mismas desde un ARNm para isoformas de peso molecular alto, así como desde sus propios ARNm añade un nuevo nivel de complejidad a la expresión del gen *EPB41*. Por otra parte, la existencia de una señal de poliadenilación en coordinación con el procesamiento alternativo que sufre el pre-ARNm de 4.1R da lugar a isoformas de 4.1R sin dominio CTD, desconocidas hasta ahora.

Todos estos datos hacen pensar que las isoformas de 4.1R son de importancia vital para la célula y que su expresión ha sido asegurada mediante diversos mecanismos.

Conclusiones



## CONCLUSIONES.

1. Los ADN codificantes para isoformas de 4.1R de peso molecular alto que contienen el exón 4 generan *in vivo* e *in vitro* dos isoformas de esta proteína, una de peso molecular alto y otra de peso molecular intermedio. Esta última, se genera por traducción interna desde el AUG2 localizado en el exón 4.
2. El mecanismo de traducción interna que permite la expresión de las isoformas de 4.1R de peso molecular intermedio, está dirigido por una secuencia de entrada interna para el ribosoma (IRES). Esta secuencia se encuentra comprendida entre el exón 2 y la secuencia anterior al AUG2 del exón 4.
3. El IRES que permite la expresión de las isoformas de 4.1R de peso molecular intermedio, está presente en dos tipos de ADN codificantes. En aquellos que codifican para isoformas de 4.1R de peso molecular alto y en aquellos que codifican únicamente para las isoformas de peso molecular intermedio.
4. La identificación de nuevos ADN codificantes para 4.1R indica la existencia de un nuevo grupo de isoformas, caracterizadas por la carencia del dominio carboxilo terminal CTD. Estas isoformas se localizan en la membrana plasmática, en el centrosoma y en el cuerpo medio de células en citocinesis.
5. Las isoformas de 4.1R sin el dominio CTD se generan a partir de un ARN mensajero en el que se introduce una señal de poliadenilación alternativa, situada en la secuencia comprendida entre el exón 17 y 18 del gen *EPB41*. Anterior a esta señal se añaden 27 nucleótidos, entre los que se incluye un codón de parada de la traducción.
6. En resumen, podemos concluir que en la generación de isoformas de 4.1R, además del procesamiento alternativo intervienen otros mecanismos como son la traducción interna y la poliadenilación alternativa. Estos datos indican que el gen *EPB41* presenta un patrón de expresión muy complejo y que las isoformas de 4.1R regulan su expresión mediante mecanismos diversos.

## Bibliografía

**Alwine, J. C., Kemp, D. J. and Stark, G. R.** (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc Natl Acad Sci U S A* **74**, 5350-4.

**Allemand, E., Guil, S., Myers, M., Moscat, J., Caceres, J. F. and Krainer, A. R.** (2005). Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proc Natl Acad Sci U S A* **102**, 3605-10.

**Anderson, R. A., Correas, I., Mazzucco, C., Castle, J. D. and Marchesi, V. T.** (1988). Tissue-specific analogues of erythrocyte protein 4.1 retain functional domains. *J Cell Biochem* **37**, 269-84.

**Anderson, R. A. and Lovrien, R. E.** (1984). Glycophorin is linked by band 4.1 protein to the human erythrocyte membrane skeleton. *Nature* **307**, 655-8.

**Ara, T., Lopez, F., Ritchie, W., Benech, P. and Gautheret, D.** (2006). Conservation of alternative polyadenylation patterns in mammalian genes. *BMC Genomics* **7**, 189.

**Aster, J. C., Welsh, M. J., Brewer, G. J. and Maisel, H.** (1984). Identification of spectrin and protein 4.1-like proteins in mammalian lens. *Biochem Biophys Res Commun* **119**, 726-34.

**Ayoubi, T. A. and Van De Ven, W. J.** (1996). Regulation of gene expression by alternative promoters. *Faseb J* **10**, 453-60.

**Baird, S. D., Turcotte, M., Korneluk, R. G. and Holcik, M.** (2006). Searching for IRES. *Rna* **12**, 1755-85.

**Baklouti, F., Huang, S. C., Tang, T. K., Delaunay, J., Marchesi, V. T. and Benz, E. J., Jr.** (1996). Asynchronous regulation of splicing events within protein 4.1 pre-mRNA during erythroid differentiation. *Blood* **87**, 3934-41.

**Baklouti, F., Huang, S. C., Vulliamy, T. J., Delaunay, J. and Benz, E. J., Jr.** (1997). Organization of the human protein 4.1 genomic locus: new insights into the tissue-specific alternative splicing of the pre-mRNA. *Genomics* **39**, 289-302.

**Baldwin, M. E., Roufail, S., Halford, M. M., Alitalo, K., Stacker, S. A. and Achen, M. G.** (2001). Multiple forms of mouse vascular endothelial growth factor-D are generated by RNA splicing and proteolysis. *J Biol Chem* **276**, 44307-14.

**Bennett, V. and Gilligan, D. M.** (1993). The spectrin-based membrane skeleton and micron-scale organization of the plasma membrane. *Annu Rev Cell Biol* **9**, 27-66.

**Blumenthal, T.** (2004). Operons in eukaryotes. *Brief Funct Genomic Proteomic* **3**, 199-211.

**Bonnal, S., Boutonnet, C., Prado-Lourenco, L. and Vagner, S.** (2003). IRESdb: the Internal Ribosome Entry Site database. *Nucleic Acids Res* **31**, 427-8.

**Bonnal, S., Pileur, F., Orsini, C., Parker, F., Pujol, F., Prats, A. C. and Vagner, S.** (2005). Heterogeneous nuclear ribonucleoprotein A1 is a novel internal ribosome entry site trans-acting factor that modulates alternative initiation of translation of the fibroblast growth factor 2 mRNA. *J Biol Chem* **280**, 4144-53.

**Burns, N. R., Ohanian, V. and Gratzer, W. B.** (1983). Properties of brain spectrin (fodrin). *FEBS Lett* **153**, 165-8.

**Calinisan, V., Gravem, D., Chen, R. P., Brittin, S., Mohandas, N., Lecomte, M. C. and Gascard, P.** (2006). New insights into potential functions for the protein 4.1 superfamily of proteins in kidney epithelium. *Front Biosci* **11**, 1646-66.

**Calkhoven, C. F., Muller, C. and Leutz, A.** (2000). Translational control of C/EBPalpha and C/EBPbeta isoform expression. *Genes Dev* **14**, 1920-32.

**Cohen, A. R., Woods, D. F., Marfatia, S. M., Walther, Z., Chishti, A. H., Anderson, J. M. and Wood, D. F.** (1998). Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *J Cell Biol* **142**, 129-38.

**Cohen, C. M., Foley, S. F. and Korsgren, C.** (1982). A protein immunologically related to erythrocyte band 4.1 is found on stress fibres on non-erythroid cells. *Nature* **299**, 648-50.

**Conboy, J. G.** (1993). Structure, function, and molecular genetics of erythroid membrane skeletal protein 4.1 in normal and abnormal red blood cells. *Semin Hematol* **30**, 58-73.

**Conboy, J. G., Chan, J. Y., Chasis, J. A., Kan, Y. W. and Mohandas, N.** (1991a). Tissue- and development-specific alternative RNA splicing regulates expression of multiple isoforms of erythroid membrane protein 4.1. *J Biol Chem* **266**, 8273-80.

**Conboy, J. G., Shitamoto, R., Parra, M., Winardi, R., Kabra, A., Smith, J. and Mohandas, N.** (1991). Hereditary elliptocytosis due to both qualitative and quantitative defects in membrane skeletal protein 4.1. *Blood* **78**, 2438-43.

**Cornelis, S., Bruynooghe, Y., Denecker, G., Van Huffel, S., Tinton, S. and Beyaert, R.** (2000). Identification and characterization of a novel cell cycle-regulated internal ribosome entry site. *Mol Cell* **5**, 597-605.

**Correas, I., Leto, T. L., Speicher, D. W. and Marchesi, V. T.** (1986a). Identification of the functional site of erythrocyte protein 4.1 involved in spectrin-actin associations. *J Biol Chem* **261**, 3310-5.

**Correas, I., Speicher, D. W. and Marchesi, V. T.** (1986b). Structure of the spectrin-actin binding site of erythrocyte protein 4.1. *J Biol Chem* **261**, 13362-6.

**Chabregas, S. M., Luche, D. D., Van Sluys, M. A., Menck, C. F. and Silva-Filho, M. C.** (2003). Differential usage of two in-frame translational start codons regulates subcellular localization of Arabidopsis thaliana THI1. *J Cell Sci* **116**, 285-91.

**Chasis, J. A., Coulombel, L., Conboy, J., McGee, S., Andrews, K., Kan, Y. W. and Mohandas, N.** (1993). Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. *J Clin Invest* **91**, 329-38.

**Chishti, A. H., Kim, A. C., Marfatia, S. M., Lutchman, M., Hanspal, M., Jindal, H., Liu, S. C., Low, P. S., Rouleau, G. A., Mohandas, N. et al.** (1998). The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. *Trends Biochem Sci* **23**, 281-2.

**Danilov, Y. N., Fennell, R., Ling, E. and Cohen, C. M.** (1990). Selective modulation of band 4.1 binding to erythrocyte membranes by protein kinase C. *J Biol Chem* **265**, 2556-62.

**Davies, G. E. and Cohen, C. M.** (1985). Platelets contain proteins immunologically related to red cell spectrin and protein 4.1. *Blood* **65**, 52-9.

**De Cárcer, G., Lallena, M. J. and Correas, I.** (1995). Protein 4.1 is a component of the nuclear matrix of mammalian cells. *Biochem J* **312** ( Pt 3), 871-7.

**De Cárcer, G., Lallena, M. J. and Correas, I.** (1995). Protein 4.1 is a component of the nuclear matrix of mammalian cells. *Biochem J* **312**, 871-7.

**Echeverri, C. J., Paschal, B. M., Vaughan, K. T. and Vallee, R. B.** (1996). Molecular characterization of the 50-kD subunit of dynactin reveals function for the complex in chromosome alignment and spindle organization during mitosis. *J Cell Biol* **132**, 617-33.

**Eder, P. S., Soong, C. J. and Tao, M.** (1986). Phosphorylation reduces the affinity of protein 4.1 for spectrin. *Biochemistry* **25**, 1764-70.

**Edwards-Gilbert, G., Veraldi, K. L. and Milcarek, C.** (1997). Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* **25**, 2547-61.

**Evan, G. I., Lewis, G. K., Ramsay, G. and Bishop, J. M.** (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol Cell Biol* **5**, 3610-6.

**Fairbanks, G., Steck, T. L. and Wallach, D. F.** (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry* **10**, 2606-17.

**Futterer, J., Kiss-Laszlo, Z. and Hohn, T.** (1993). Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **73**, 789-802.

**Gilbert, W.** (1978). Why genes in pieces? *Nature* **271**, 501.

**Giraud, S., Greco, A., Brink, M., Diaz, J. J. and Delafontaine, P.** (2001). Translation initiation of the insulin-like growth factor I receptor mRNA is mediated by an internal ribosome entry site. *J Biol Chem* **276**, 5668-75.

**Goodman, S. R., Casoria, L. A., Coleman, D. B. and Zagon, I. S.** (1984). Identification and location of brain protein 4.1. *Science* **224**, 1433-6.

**Granger, B. L. and Lazarides, E.** (1984). Membrane skeletal protein 4.1 of avian erythrocytes is composed of multiple variants that exhibit tissue-specific expression. *Cell* **37**, 595-607.

**Granger, B. L. and Lazarides, E.** (1985). Appearance of new variants of membrane skeletal protein 4.1 during terminal differentiation of avian erythroid and lenticular cells. *Nature* **313**, 238-41.

**Gunning, P. W., Schevzov, G., Kee, A. J. and Hardeman, E. C.** (2005). Tropomyosin isoforms: divining rods for actin cytoskeleton function. *Trends Cell Biol* **15**, 333-41.

**Han, B. and Zhang, J. T.** (2002). Regulation of gene expression by internal ribosome entry sites or cryptic promoters: the eIF4G story. *Mol Cell Biol* **22**, 7372-84.

**Han, B. G., Nunomura, W., Takakuwa, Y., Mohandas, N. and Jap, B. K.** (2000). Protein 4.1R core domain structure and insights into regulation of cytoskeletal organization. *Nat Struct Biol* **7**, 871-5.

**Hanahan, D.** (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-80.

**Hattori, M.** (2005). [Finishing the euchromatic sequence of the human genome]. *Tanpakushitsu Kakusan Koso* **50**, 162-8.

**Henis-Korenblit, S., Strumpf, N. L., Goldstaub, D. and Kimchi, A.** (2000). A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation. *Mol Cell Biol* **20**, 496-506.

**Hochheimer, A. and Tjian, R.** (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes Dev* **17**, 1309-20.

**Holcik, M., Gordon, B. W. and Korneluk, R. G.** (2003). The internal ribosome entry site-mediated translation of antiapoptotic protein XIAP is modulated by the heterogeneous nuclear ribonucleoproteins C1 and C2. *Mol Cell Biol* **23**, 280-8.

**Holcik, M. and Korneluk, R. G.** (2000). Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: role of La autoantigen in XIAP translation. *Mol Cell Biol* **20**, 4648-57.

**Holcik, M., Lefebvre, C., Yeh, C., Chow, T. and Korneluk, R. G.** (1999). A new internal-ribosome-entry-site motif potentiates XIAP-mediated cytoprotection. *Nat Cell Biol* **1**, 190-2.

**Hoover, K. B. and Bryant, P. J.** (2000). The genetics of the protein 4.1 family: organizers of the membrane and cytoskeleton. *Curr Opin Cell Biol* **12**, 229-34.

**Hou, C. L., Tang, C., Roffler, S. R. and Tang, T. K.** (2000). Protein 4.1R binding to eIF3-p44 suggests an interaction between the cytoskeletal network and the translation apparatus. *Blood* **96**, 747-53.

**Jackson, R. J.** (2005). Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem Soc Trans* **33**, 1231-41.

**Jons, T. and Drenckhahn, D.** (1992). Identification of the binding interface involved in linkage of cytoskeletal protein 4.1 to the erythrocyte anion exchanger. *Embo J* **11**, 2863-7.

**Kelly, G. M., Zelus, B. D. and Moon, R. T.** (1991). Identification of a calcium-dependent calmodulin-binding domain in *Xenopus* membrane skeleton protein 4.1. *J Biol Chem* **266**, 12469-73.

**Kontogianni-Konstantopoulos, A., Frye, C. S., Benz, E. J., Jr. and Huang, S. C.** (2001a). The prototypical 4.1R-10-kDa domain and the 4.1g-10-kDa paralog mediate fodrin-actin complex formation. *J Biol Chem* **276**, 20679-87.

**Kontogianni-Konstantopoulos, A., Frye, C. S., Benz, E. J., Jr. and Huang, S. C.** (2001b). The prototypical 4.1R-10-kDa domain and the 4.1g-10-kDa paralog mediate fodrin-actin complex formation. *J Biol Chem* **276**, 20679-87.

**Kontogianni-Konstantopoulos, A., Huang, S. C. and Benz, E. J., Jr.** (2000). A nonerythroid isoform of protein 4.1R interacts with components of the contractile apparatus in skeletal myofibers. *Mol Biol Cell* **11**, 3805-17.

**Koromilas, A. E., Lazaris-Karatzas, A. and Sonenberg, N.** (1992). mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *Embo J* **11**, 4153-8.

**Kozak, M.** (2001). New ways of initiating translation in eukaryotes? *Mol Cell Biol* **21**, 1899-907.

**Kozak, M.** (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**, 1-34.

**Kozak, M.** (2003). Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation. *Gene* **318**, 1-23.

**Krauss, S. W., Larabell, C. A., Lockett, S., Gascard, P., Penman, S., Mohandas, N. and Chasis, J. A.** (1997). Structural protein 4.1 in the nucleus of human cells: dynamic rearrangements during cell division. *J Cell Biol* **137**, 275-89.

**Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

**Lallena, M. J. and Correias, I.** (1997). Transcription-dependent redistribution of nuclear protein 4.1 to SC35- enriched nuclear domains. *J Cell Sci* **110**, 239-47.

**Lallena, M. J., Martínez, C., Valcárcel, J. and Correias, I.** (1998). Functional association of nuclear protein 4.1 with pre-mRNA splicing factors. *J Cell Sci* **111**, 1963-71.

**Lambert, S., Conboy, J. and Zail, S.** (1988). A molecular study of heterozygous protein 4.1 deficiency in hereditary elliptocytosis. *Blood* **72**, 1926-9.

**Leclerc, E. and Vetter, S.** (1998). Characterization of a calcium-dependent calmodulin-binding domain in the 135-kD human protein 4.1 isoform. *Eur J Biochem* **258**, 567-71.



**Lee, S. J.** (1991). Expression of growth/differentiation factor 1 in the nervous system: conservation of a bicistronic structure. *Proc Natl Acad Sci U S A* **88**, 4250-4.

**Leto, T. L., Pratt, B. M. and Madri, J. A.** (1986). Mechanisms of cytoskeletal regulation: modulation of aortic endothelial cell protein band 4.1 by the extracellular matrix. *J Cell Physiol* **127**, 423-31.

**Ling, E., Danilov, Y. N. and Cohen, C. M.** (1988). Modulation of red cell band 4.1 function by cAMP-dependent kinase and protein kinase C phosphorylation. *J Biol Chem* **263**, 2209-16.

**Liu, Z., Dong, Z., Han, B., Yang, Y., Liu, Y. and Zhang, J. T.** (2005). Regulation of expression by promoters versus internal ribosome entry site in the 5'-untranslated sequence of the human cyclin-dependent kinase inhibitor p27kip1. *Nucleic Acids Res* **33**, 3763-71.

**Lou, H., Yang, Y., Cote, G. J., Berget, S. M. and Gagel, R. F.** (1995). An intron enhancer containing a 5' splice site sequence in the human calcitonin/calcitonin gene-related peptide gene. *Mol Cell Biol* **15**, 7135-42.

**Lue, R. A., Marfatia, S. M., Branton, D. and Chishti, A. H.** (1994). Cloning and characterization of hdlg: the human homologue of the Drosophila discs large tumor suppressor binds to protein 4.1. *Proc Natl Acad Sci U S A* **91**, 9818-22.

**Luque, C. M. and Correas, I.** (2000). A constitutive region is responsible for nuclear targeting of 4.1R: modulation by alternative sequences results in differential intracellular localization. *J Cell Sci* **113**, 2485-2495.

**Luque, C. M., Lallena, M. J., Pérez-Ferreiro, C. M., de Isidro, Y., De Cárcer, G., Alonso, M. A. and Correas, I.** (1999). The N-terminal 209-aa domain of high molecular-weight 4.1R isoforms abrogates 4.1R targeting to the nucleus. *Proc Natl Acad Sci U S A* **96**, 14925-30.

**Maier, D., Nagel, A. C. and Preiss, A.** (2002). Two isoforms of the Notch antagonist Hairless are produced by differential translation initiation. *Proc Natl Acad Sci U S A* **99**, 15480-5.

**Mangus, D. A., Evans, M. C. and Jacobson, A.** (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* **4**, 223.

**Maniatis, T. and Reed, R.** (2002). An extensive network of coupling among gene expression machines. *Nature* **416**, 499-506.

**Maniatis, T. and Tasic, B.** (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **418**, 236-43.

**Marfatia, S. M., Leu, R. A., Branton, D. and Chishti, A. H.** (1995). Identification of the protein 4.1 binding interface on glycophorin C and p55, a homologue of the Drosophila discs-large tumor suppressor protein. *J Biol Chem* **270**, 715-9.

**Marfatia, S. M., Lue, R. A., Branton, D. and Chishti, A. H.** (1994). In vitro binding studies suggest a membrane-associated complex between erythroid p55, protein 4.1, and glycophorin C. *J Biol Chem* **269**, 8631-4.

**Mattagajasingh, S. N., Huang, S. C., Hartenstein, J. S. and Benz, E. J., Jr.** (2000). Characterization of the interaction between protein 4.1R and ZO-2. A possible link between the tight junction and the actin cytoskeleton. *J Biol Chem* **275**, 30573-85.

**Mattagajasingh, S. N., Huang, S. C., Hartenstein, J. S., Snyder, M., Marchesi, V. T. and Benz, E. J.** (1999). A nonerythroid isoform of protein 4.1R interacts with the nuclear mitotic apparatus (NuMA) protein. *J Cell Biol* **145**, 29-43.



**Meyer, E. L., Strutz, N., Gahring, L. C. and Rogers, S. W.** (2003). Glutamate receptor subunit 3 is modified by site-specific limited proteolysis including cleavage by gamma-secretase. *J Biol Chem* **278**, 23786-96.

**Modrek, B. and Lee, C.** (2002). A genomic view of alternative splicing. *Nat Genet* **30**, 13-9.

**Morrish, B. C. and Rumsby, M. G.** (2002). The 5' untranslated region of protein kinase Cdelta directs translation by an internal ribosome entry segment that is most active in densely growing cells and during apoptosis. *Mol Cell Biol* **22**, 6089-99.

**Norton, P. A.** (1994). Alternative pre-mRNA splicing: factors involved in splice site selection. *J Cell Sci* **107** ( Pt 1), 1-7.

**Parra, M., Gascard, P., Walensky, L. D., Gimm, J. A., Blackshaw, S., Chan, N., Takakuwa, Y., Berger, T., Lee, G., Chasis, J. A. et al.** (2000). Molecular and functional characterization of protein 4.1B, a novel member of the protein 4.1 family with high level, focal expression in brain. *J Biol Chem* **275**, 3247-55.

**Parra, M., Gascard, P., Walensky, L. D., Snyder, S. H., Mohandas, N. and Conboy, J. G.** (1998). Cloning and characterization of 4.1G (EPB41L2), a new member of the skeletal protein 4.1 (EPB41) gene family. *Genomics* **49**, 298-306.

**Parra, M. K., Gee, S. L., Koury, M. J., Mohandas, N. and Conboy, J. G.** (2003). Alternative 5' exons and differential splicing regulate expression of protein 4.1R isoforms with distinct N-termini. *Blood* **101**, 4164-71.

**Pasternack, G. R., Anderson, R. A., Leto, T. L. and Marchesi, V. T.** (1985). Interactions between protein 4.1 and band 3. An alternative binding site for an element of the membrane skeleton. *J Biol Chem* **260**, 3676-83.

**Pasternack, G. R. and Racusen, R. H.** (1989). Erythrocyte protein 4.1 binds and regulates myosin. *Proc Natl Acad Sci U S A* **86**, 9712-6.

**Perez-Ferreiro, C. M., Lospitao, E. and Correas, I.** (2006). Protein 4.1R self-association: identification of the binding domain. *Biochem J* **400**, 457-65.

**Perez-Ferreiro, C. M., Luque, C. M. and Correas, I.** (2001). 4.1R proteins associate with interphase microtubules in human T cells: a 4.1R constitutive region is involved in tubulin binding. *J Biol Chem* **276**, 44785-91.

**Perez-Ferreiro, C. M., Vernos, I. and Correas, I.** (2004). Protein 4.1R regulates interphase microtubule organization at the centrosome. *J Cell Sci* **117**, 6197-206.

**Peters, L. L., Weier, H. U., Walensky, L. D., Snyder, S. H., Parra, M., Mohandas, N. and Conboy, J. G.** (1998). Four paralogous protein 4.1 genes map to distinct chromosomes in mouse and human. *Genomics* **54**, 348-50.

**Pickering, B. M., Mitchell, S. A., Evans, J. R. and Willis, A. E.** (2003). Polypyrimidine tract binding protein and poly r(C) binding protein 1 interact with the BAG-1 IRES and stimulate its activity in vitro and in vivo. *Nucleic Acids Res* **31**, 639-46.

**Pilipenko, E. V., Pestova, T. V., Kolupaeva, V. G., Khitrina, E. V., Poperechnaya, A. N., Agol, V. I. and Hellen, C. U.** (2000). A cell cycle-dependent protein serves as a template-specific translation initiation factor. *Genes Dev* **14**, 2028-45.

**Pyronnet, S., Pradayrol, L. and Sonenberg, N.** (2000). A cell cycle-dependent internal ribosome entry site. *Mol Cell* **5**, 607-16.

**Pyronnet, S. and Sonenberg, N.** (2001). Cell-cycle-dependent translational control. *Curr Opin Genet Dev* **11**, 13-8.

**Sambrook, J., Fritsch, E. F. and Maniatis, T.** (1989). Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

**Schaub, M. and Keller, W.** (2002). RNA editing by adenosine deaminases generates RNA and protein diversity. *Biochimie* **84**, 791-803.

**Schischmanoff, P. O., Winardi, R., Discher, D. E., Parra, M. K., Bicknese, S. E., Witkowska, H. E., Conboy, J. G. and Mohandas, N.** (1995). Defining of the minimal domain of protein 4.1 involved in spectrin-actin binding. *J Biol Chem* **270**, 21243-50.

**Schischmanoff, P. O., Yaswen, P., Parra, M. K., Lee, G., Chasis, J. A., Mohandas, N. and Conboy, J. G.** (1997). Cell shape-dependent regulation of protein 4.1 alternative pre-mRNA splicing in mammary epithelial cells. *J Biol Chem* **272**, 10254-9.

**Shi, Z. T., Afzal, V., Coller, B., Patel, D., Chasis, J. A., Parra, M., Lee, G., Paszty, C., Stevens, M., Walensky, L. et al.** (1999). Protein 4.1R-deficient mice are viable but have erythroid membrane skeleton abnormalities. *J Clin Invest* **103**, 331-40.

**Shin, D., Park, S. and Park, C.** (2003). A splice variant acquiring an extra transcript leader region decreases the translation of glutamine synthetase gene. *Biochem J* **374**, 175-84.

**Smith, C. W. and Valcarcel, J.** (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci* **25**, 381-8.

**Sonenberg, N. and Dever, T. E.** (2003). Eukaryotic translation initiation factors and regulators. *Curr Opin Struct Biol* **13**, 56-63.

**Southan, C.** (2004). Has the yo-yo stopped? An assessment of human protein-coding gene number. *Proteomics* **4**, 1712-26.

**Spiegel, J. E., Beardsley, D. S., Southwick, F. S. and Lux, S. E.** (1984). An analogue of the erythroid membrane skeletal protein 4.1 in nonerythroid cells. *J Cell Biol* **99**, 886-93.

**Steck, T. L.** (1974). The organization of proteins in the human red blood cell membrane. A review. *J Cell Biol* **62**, 1-19.

**Stein, I., Itin, A., Einat, P., Skaliter, R., Grossman, Z. and Keshet, E.** (1998). Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia. *Mol Cell Biol* **18**, 3112-9.

**Stoneley, M., Chappell, S. A., Jopling, C. L., Dickens, M., MacFarlane, M. and Willis, A. E.** (2000). c-Myc protein synthesis is initiated from the internal ribosome entry segment during apoptosis. *Mol Cell Biol* **20**, 1162-9.

**Strachan, T. a. R., Andrew P.** (2003). Human Molecular Genetics 2.

**Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F. et al.** (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A* **99**, 16899-903.

**Strehler, E. E. and Zacharias, D. A.** (2001). Role of alternative splicing in generating isoform diversity among plasma membrane calcium pumps. *Physiol Rev* **81**, 21-50.

**Subrahmanyam, G., Bertics, P. J. and Anderson, R. A.** (1991). Phosphorylation of protein 4.1 on tyrosine-418 modulates its function in vitro. *Proc Natl Acad Sci U S A* **88**, 5222-6.

**Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M. and Arai, N.** (1988). SR $\alpha$  promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cell Biol* **8**, 466-72.

- Tanaka, T., Kadowaki, K., Lazarides, E. and Sobue, K.** (1991).  $\text{Ca}^{2+}$ -dependent regulation of the spectrin/actin interaction by calmodulin and protein 4.1. *J Biol Chem* **266**, 1134-40.
- Tang, T. K., Leto, T. L., Correas, I., Alonso, M. A., Marchesi, V. T. and Benz, E. J., Jr.** (1988). Selective expression of an erythroid-specific isoform of protein 4.1. *Proc Natl Acad Sci U S A* **85**, 3713-7.
- Tang, T. K., Qin, Z., Leto, T., Marchesi, V. T. and Benz, E. J., Jr.** (1990). Heterogeneity of mRNA and protein products arising from the protein 4.1 gene in erythroid and nonerythroid tissues. *J Cell Biol* **110**, 617-24.
- Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A. C. and Vagner, S.** (2003). Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. *Biol Cell* **95**, 169-78.
- Tyler, J. M., Reinhardt, B. N. and Branton, D.** (1980). Associations of erythrocyte membrane proteins. Binding of purified bands 2.1 and 4.1 to spectrin. *J Biol Chem* **255**, 7034-9.
- Vagner, S., Galy, B. and Pyronnet, S.** (2001). Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. *EMBO Rep* **2**, 893-8.
- Vattem, K. M. and Wek, R. C.** (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A* **101**, 11269-74.
- Venezia, N. D., Maillet, P., Morle, L., Roda, L., Delaunay, J. and Baklouti, F.** (1998). A large deletion within the protein 4.1 gene associated with a stable truncated mRNA and an unaltered tissue-specific alternative splicing. *Blood* **91**, 4361-7.
- Vetter, S. W. and Leclerc, E.** (2001). Phosphorylation of serine residues affects the conformation of the calmodulin binding domain of human protein 4.1. *Eur J Biochem* **268**, 4292-9.
- Wahle, E. and Ruegsegger, U.** (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev* **23**, 277-95.
- Walensky, L. D., Blackshaw, S., Liao, D., Watkins, C. C., Weier, H. U., Parra, M., Haganir, R. L., Conboy, J. G., Mohandas, N. and Snyder, S. H.** (1999). A novel neuron-enriched homolog of the erythrocyte membrane cytoskeletal protein 4.1. *J Neurosci* **19**, 6457-67.
- Walensky, L. D., Shi, Z. T., Blackshaw, S., DeVries, A. C., Demas, G. E., Gascard, P., Nelson, R. J., Conboy, J. G., Rubin, E. M., Snyder, S. H. et al.** (1998). Neurobehavioral deficits in mice lacking the erythrocyte membrane cytoskeletal protein 4.1. *Curr Biol* **8**, 1269-72.
- Wedekind, J. E., Dance, G. S., Sowden, M. P. and Smith, H. C.** (2003). Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet* **19**, 207-16.
- Woodley, L. and Valcarcel, J.** (2002). Regulation of alternative pre-mRNA splicing. *Brief Funct Genomic Proteomic* **1**, 266-77.
- Yamakawa, H., Ohara, R., Nakajima, D., Nakayama, M. and Ohara, O.** (1999). Molecular characterization of a new member of the protein 4.1 family (brain 4.1) in rat brain. *Brain Res Mol Brain Res* **70**, 197-209.

Anexo

# Protein 4.1R self-association: identification of the binding domain

Carmen M. PÉREZ-FERREIRO<sup>1</sup>, Eva LOSPITAO and Isabel CORREAS

Departamento de Biología Molecular, CBMSO (Centro de Biología Molecular Severo Ochoa), UAM/CSIC (Universidad Autónoma de Madrid, Consejo Superior de Investigaciones Científicas), Cantoblanco, E-28049 Madrid, Spain

Erythroid protein 4.1 (4.1R) stabilizes the spectrin–actin network and anchors it to the plasma membrane. To contribute to the characterization of non-erythroid protein 4.1R, we used sedimentation, pull-down and co-immunoprecipitation assays to investigate the ability of protein 4.1R to establish inter-/intra-molecular associations. We demonstrated that the small 4.1R isoforms of 60 kDa (4.1R<sup>60</sup>), but not the larger isoforms of 80 and 135 kDa (4.1R<sup>80</sup> and 4.1R<sup>135</sup>), were self-associated, and that a domain contained in all 4.1R isoforms, the core region, was responsible for 4.1R self-association. Results from denaturing–renaturing experiments, in which an initially non-self-associated 4.1R<sup>80</sup> isoform became self-associated, suggested that an initially hidden core region was subsequently exposed. This hypothesis was sup-

ported by results from pull-down assays, which showed that the core region interacted with the N-terminal end of the FERM (4.1, ezrin, radixin, moesin) domain that is present in 4.1R<sup>80</sup> and 4.1R<sup>135</sup> isoforms but absent from 4.1R<sup>60</sup> isoforms. Consistently, 4.1R<sup>80</sup> isoforms bound neither to each other nor to 4.1R<sup>60</sup> isoforms. We propose that 4.1R<sup>60</sup> isoforms are constitutively self-associated, whereas 4.1R<sup>80</sup> and 4.1R<sup>135</sup> self-association is prevented by intramolecular interactions.

**Key words:** core region, FERM (4.1, ezrin, radixin, moesin) domain, intermolecular interaction, 4.1R isoform, self-association.

## INTRODUCTION

Protein 4.1R was first identified in human erythrocytes as an 80 kDa multifunctional protein of the membrane skeleton. Erythroid protein 4.1R is essential for maintaining erythrocyte shape and mechanical properties of the membrane, such as deformability and stability. In this cell type, protein 4.1R stabilizes the spectrin–actin network and mediates the attachment of the underlying cytoskeleton to the overlaying lipid bilayer through interactions with lipids (phosphatidylserine and phosphatidylinositol 4,5-bisphosphate) [1,2] as well as with integral membrane proteins [3]. Deficiency of 4.1R in erythrocytes leads to the assembly of an unstable cytoskeleton structure that manifests itself as hereditary elliptocytosis, a disease characterized by the loss of normal discoid morphology and the presence of oval or elliptical erythrocytes with unstable membranes [3].

Subsequent studies showed that multiple immunoreactive 4.1R proteins, varying in size from 30 to 210 kDa, are present in nucleated cells [4,5]. Isoforms of 4.1R are expressed in many tissues and detected at different subcellular locations [6–9]. The roles and partners of 4.1R in non-erythroid cells are beginning to be elucidated [10–21].

The multiple isoforms of 4.1R are mainly expressed as a result of extensive alternative splicing of the 4.1R-encoding pre-mRNA [22,23]. This event is cell- and tissue-specific and dependent on the growth and differentiation stages of the cell [24–30]. The prototypical erythroid protein 4.1R<sup>80</sup> is produced when 17 nt 5'-upstream from exon 2 are spliced out, and translation is initiated at the downstream start site present in exon 4 (ATG2). Larger isoforms, termed 4.1R<sup>135</sup>, containing up to 209 amino acids at the N-terminus of the erythroid 4.1R<sup>80</sup>, are synthesized when the 17-nt sequence containing the upstream ATG (ATG1) translation initiation codon is included. These isoforms are predominantly expressed in non-erythroid cells. A third type of isoforms, termed

4.1R<sup>60</sup>, can be produced in erythroid and non-erythroid cells when both the 17-nt sequence (containing the ATG1) and exon 4 (containing the ATG2) are spliced out and translation is initiated from a third translation-initiation codon (ATG3) present in exon 8 [31,32]. Thus three different types of 4.1R isoforms varying in their N-terminal extensions can be generated depending on whether ATG1, ATG2 or ATG3 is used as the translation-initiation codon.

Erythroid 4.1R<sup>80</sup> protein contains four structural domains [33]: a 30 kDa N-terminal membrane-binding domain that has recently been named the FERM domain (4.1, ezrin, radixin, moesin) [34,35], a 16 kDa domain, a 10 kDa SAB (spectrin/actin-binding domain) and the 22–24-kDa CTD (C-terminal domain). 4.1R<sup>135</sup> isoforms contain an extra N-terminal extension domain [HP (Head-Piece)] comprised of up to 209 amino acids, whereas 4.1R<sup>60</sup> isoforms lack the HP domain and the N-terminal end of the FERM domain.

In recent years our group has been interested in characterizing non-erythroid protein 4.1R. Systematic studies using 4.1R<sup>135</sup>, 4.1R<sup>80</sup> and 4.1R<sup>60</sup> isoform types showed that the small 4.1R<sup>60</sup> isoforms present significant differences in their subcellular distribution and function [15,32,36]. To gain further insights into non-erythroid protein 4.1R, in the present study we have analysed the biochemical behaviour of different 4.1R isoform types using sedimentation, co-immunoprecipitation and pull-down assays. These experiments show that 4.1R<sup>60</sup>, but not 4.1R<sup>135</sup> or 4.1R<sup>80</sup> isoforms, establishes intermolecular interactions through a constitutive domain, termed the core region. Since this region is present in all 4.1R isoforms and only 4.1R<sup>60</sup> isoforms were found to be self-associated, we reasoned that 4.1R<sup>135</sup> and 4.1R<sup>80</sup> isoforms should also have the ability to self-associate, but that their core region needs to be exposed for such an interaction to occur. Results from subsequent experiments are consistent with this idea. Our results indicate that all 4.1R proteins have the ability

Abbreviations used: Cter, calmodulin, troponin C, essential and regulatory light chains of myosin; ERM, ezrin, radixin, moesin; FERM, 4.1, ezrin, radixin, moesin; GFP, green fluorescent protein; GST, glutathione S-transferase.

<sup>1</sup> To whom correspondence should be addressed (email cmperez@cbm.uam.es).

to self-associate but that large 4.1R<sup>135</sup> and 4.1R<sup>80</sup> isoforms prevent their self-association by intramolecular interaction between the N-terminus of the FERM domain and the self-association region.

## EXPERIMENTAL

### Cell culture and transfection

COS-7 cells were grown as described in [14]. Transfection experiments were performed by electroporation using the Electro Cell Manipulator 600 (BTX, San Diego, CA, U.S.A.). Cells were processed 48 h after transfection.

### cDNA cloning and composite cDNA constructs

4.1R<sup>135</sup>Δ16; 4.1R<sup>80</sup>Δ16; 4.1R<sup>60</sup>Δ16 and 4.1R<sup>60</sup>Δ16,18 cDNAs were cloned from Molt-4 T-cells and tagged as described in [32,36]. 4.1R<sup>135</sup>Δ16-GFP (green fluorescent protein); 4.1R<sup>80</sup>Δ16-GFP; and 4.1R<sup>60</sup>Δ16,18-GFP were constructed as detailed in [37]. GST (glutathione S-transferase); GST-4.1R<sup>135</sup>Δ16; GST-4.1R<sup>80</sup>Δ16; GST-4.1R<sup>60</sup>Δ16,18; GST-Cter (calmodulin, troponin C, essential and regulatory light chains of myosin); GST-core and GST-coreΔLeu proteins were prepared as described in [14]. GST-4.1R<sup>E4-E8</sup> was constructed by PCR using pSRα4.1R<sup>80</sup>Δ16 as a template [32]. Appropriate sense and antisense primers containing the BglII and XhoI restriction sites at the 5'- and 3'-ends respectively were used for the amplification reactions. The amplified cDNAs were inserted into the BamHI and XhoI sites of pGEX-6P1 vector (Amersham Biosciences, Piscataway, NJ, U.S.A.) in-frame with the GST coding sequence. GST fusion proteins were cleaved by PreScission protease (Amersham Biosciences) following the manufacturer's instructions.

### Antibodies

Anti-4.1R (10b) antibody is an affinity-purified polyclonal antibody generated as described previously [38] and recognizes a sequence encoded by exon 17. Anti-4.1R (764) is a polyclonal antibody raised against a synthetic peptide (FRYSGRTQAQTRC) whose sequence is encoded by exon 12. Anti-GST and anti-FLAG antibodies are rabbit polyclonal antibodies (Sigma, St. Louis, MO, U.S.A.). Anti-GFP is a rabbit polyclonal antibody (Molecular Probes, Eugene, OR, U.S.A.). Horseradish-peroxidase-labelled secondary antibodies were obtained from Southern Biotechnology Associates.

### Western blot analysis and immunoprecipitation assays

Protein samples were separated by SDS/PAGE and transferred to Immobilon (Millipore) in Tris/borate buffer (pH 8.2). Membranes were processed and developed as described in [18]. For immunoprecipitation assays, COS-7 cells were washed twice with PBS and scraped from the plate into PBS containing 5 mM EDTA. Cells were centrifuged, and the pellet was resuspended in lysis buffer [10 mM Tris/HCl (pH 7.6) 150 mM NaCl, 1 % Nonidet P40, 5 mM EDTA, 0.5 mM PMSF and 1 μg/ml each of leupeptin, aprotinin and pepstatin], incubated on ice for 20 min and centrifuged in a Minifuge for 10 min at 4°C. The supernatant was incubated with antibody-coupled Sepharose beads for 3 h at 4°C and processed as previously described [39].

### Sucrose-density-gradient centrifugation assays

COS-7 cells were harvested, lysed and sedimented as described in [40]. Briefly, cells were lysed in a buffer containing 50 mM Tris/HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 % Nonidet

P40, 0.5 mM PMSF and 1 μg/ml each of leupeptin, aprotinin and pepstatin, and kept at 4°C for 15 min. The lysates were centrifuged at 100 000 *g* for 25 min at 4°C in a Beckman TL-100 table-top ultracentrifuge using a TLA-100.1 fixed-angle rotor. The supernatants were loaded on to 4.2 ml linear 5–20 % (w/v) sucrose gradients prepared in lysis buffer without Nonidet P40. Gradients were centrifuged in a TST-60.4 rotor (Kontron Instruments) at 26 000 rev./min for 18 h at 4°C and collected as 300 μl fractions. Sedimentation standards BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S) were run in parallel. Equal volumes of each fraction were analysed by SDS/PAGE followed by Western blotting [18]. For protein denaturation and renaturation (Figure 5), GST-4.1R<sup>80</sup>Δ16 protein was dialysed serially against buffer A (50 mM Tris/HCl, pH 7.5, 150 mM NaCl and 1 mM EDTA) containing 2, 4 or 8 M urea. The protein was slowly renatured by removing the urea by serially dialysing against buffer A containing 4, 2 and 0 M urea. After completely removing the urea, the GST-4.1R<sup>80</sup>Δ16 protein was loaded on to the sucrose gradients, centrifuged, fractionated and processed as described above. For recombinant proteins 10 μg of fusion protein was loaded on to the sucrose gradient.

### In vitro protein expression

*In vitro* protein expression was achieved by coupled *in vitro* transcription and translation reactions using the TNT T7 reticulocyte lysate system (Promega, Madison, WI, U.S.A.) as previously described [41]. Synthesized proteins were radiolabelled by including [<sup>35</sup>S]methionine (<sup>35</sup>S-Met; Amersham Biosciences) in the reaction.

### Pull-down assays

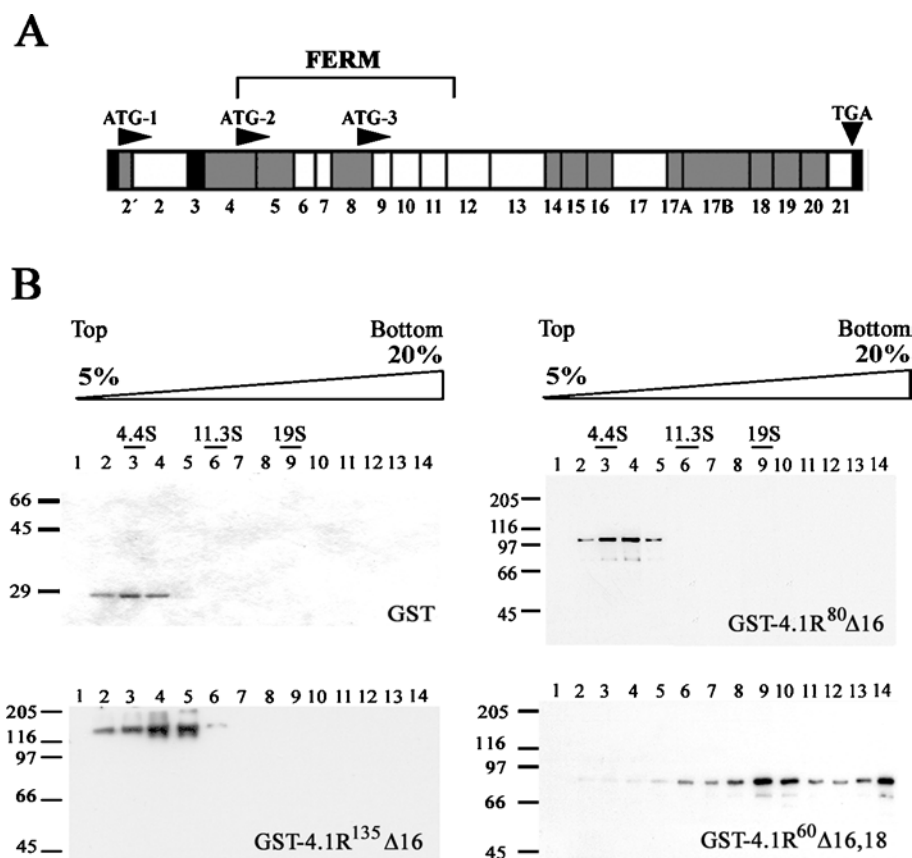
GST and the recombinant proteins GST-4.1R<sup>80</sup>Δ16, GST-4.1R<sup>60</sup>Δ16,18, GST-Cter, GST-core, GST-coreΔLeu and GST-4.1R<sup>E4-E8</sup> were prepared as indicated in [14]. COS-7 cells were lysed and processed as described in [15]. Briefly, COS-7 cell lysates or radiolabelled proteins were incubated for 1 h at 4°C with the glutathione-Sepharose-4B column loaded with the corresponding GST proteins. After extensive washes the beads were resuspended in Laemmli buffer and boiled for 5 min. The bound proteins were separated by SDS/PAGE, transferred to Immobilon membranes according to standard procedures and visualized by immunoblotting as described above.

### Gel filtration

Gel filtration chromatography was carried out using a Superdex 200 HiLoad 16/60 column (Amersham Biosciences) controlled by an AKTA FPLC system (Amersham Biosciences). The column was equilibrated with two column volumes of a buffer containing 50 mM Tris (pH 7.5), 150 mM NaCl and 1 mM EDTA and run at 4°C. Globular proteins of known molecular mass were used to calibrate the column before applying protein GST-4.1R<sup>60</sup>Δ16, 18. The marker proteins used were thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), BSA (66 kDa), ovalbumin (43 kDa) and chymotrypsinogen A (25 kDa). Proteins were eluted at a flow rate of 0.4 ml/min. The eluate was monitored by absorbance at 280 nm and the collected fractions (1.6 ml/fraction) were analysed by SDS/PAGE. The distribution of protein GST-4.1R<sup>60</sup>Δ16,18 was detected with the 10b anti-4.1R antibody.

### Sedimentation velocity analysis

Analytical ultracentrifugation experiments were carried out at 45 000 rev./min and 11.5°C for 3 h in an XL-A analytical



**Figure 1** A fusion protein containing the small 4.1R<sup>60</sup>Δ16,18 isoform migrates with a high sedimentation coefficient in sedimentation assays

(A) Schematic representation of the exon map for the 4.1R protein. Exons are coded as follows: grey, alternative; white, constitutive; black, non-coding. The number of each exon is shown at the bottom of the scheme. Three translation-initiation sites at exons 2' (ATG-1), 4 (ATG-2) and 8 (ATG-3) and the stop codon (TGA) at exon 21 are shown. (B) GST and the fusion proteins GST-4.1R<sup>135</sup>Δ16 and GST-4.1R<sup>60</sup>Δ16,18 were purified and loaded on to 5–20% sucrose-density gradients, centrifuged and fractionated as described in the Experimental section. Fractions were analysed by immunoblotting using the anti-4.1R antibody (results not shown). The sucrose gradient, the fractions collected and the position of sedimentation of the standard proteins, BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S), are all indicated. Molecular size markers are reported on the left, in kilodaltons. Note that the smallest protein, GST-4.1R<sup>60</sup>Δ16,18, migrates with a higher sedimentation coefficient than that of protein GST-4.1R<sup>80</sup>Δ16 and GST-4.1R<sup>135</sup>Δ16.

ultracentrifuge (Beckman Coulter) equipped with UV-visible absorbance optics, an An50Ti rotor and 12 mm double-sector centrepieces. The protein (loading concentration of 0.5 mg/ml) was equilibrated in 50 mM Tris (pH 7.5), 150 mM NaCl and 1 mM EDTA. Data were collected at 280 nm. Differential sedimentation coefficient distributions,  $c(s)$ , were calculated by least-squares boundary modelling of sedimentation velocity data using the program SEDFIT [42].

## RESULTS

### Isoforms of protein 4.1R show different sedimentation behaviours

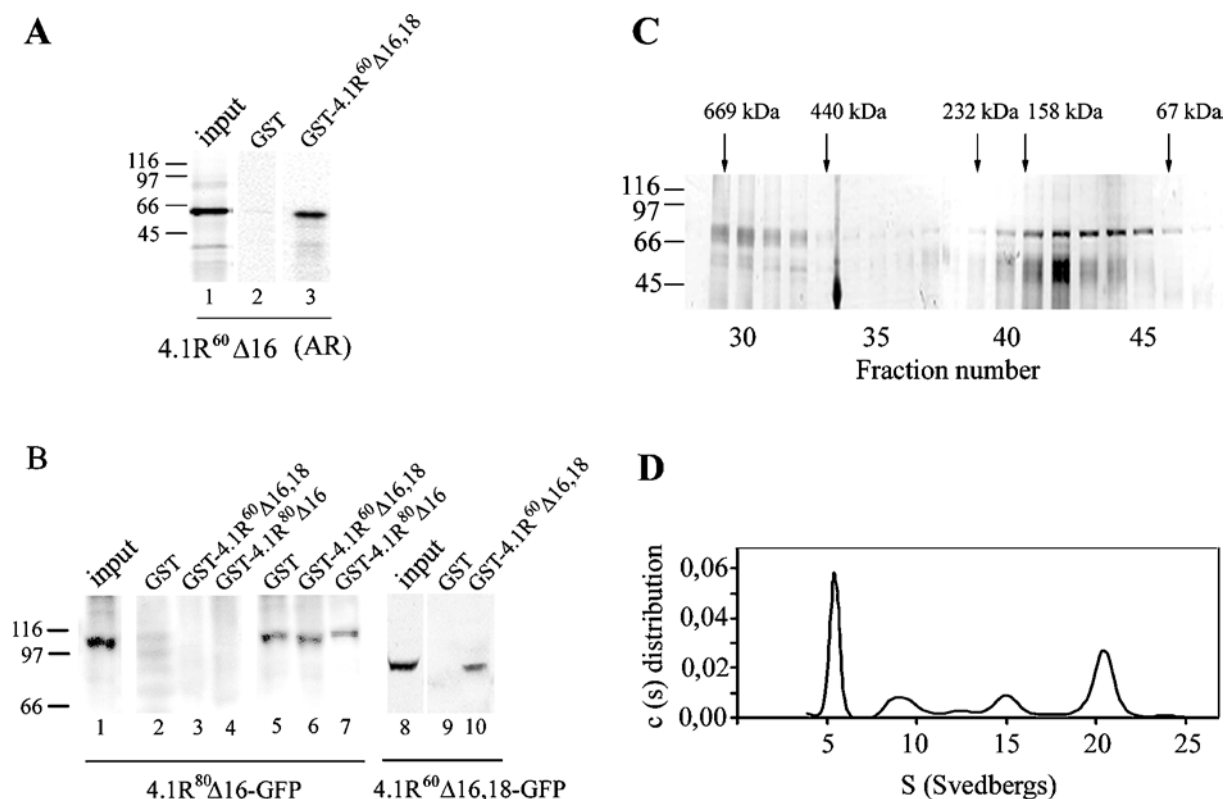
To gain insights into non-erythroid protein 4.1R, we analysed the biochemical behaviour of different 4.1R isoform types using sucrose-density-gradient centrifugation assays. We assayed 4.1R isoforms translated from the most upstream start codon, ATG1 (4.1R<sup>135</sup> isoforms; ~135 kDa), and from the downstream start sites, ATG2 (4.1R<sup>80</sup> isoforms; ~80 kDa) and ATG3 (4.1R<sup>60</sup> isoforms; ~60 kDa) (Figure 1A) fused to GST. We chose representative isoforms that correspond to the most abundant of each type of 4.1R isoforms in non-erythroid human T-cells and whose cDNAs we had previously isolated [32,36]. Fusion proteins comprised of GST and full-length 4.1R<sup>135</sup>Δ16, 4.1R<sup>80</sup>Δ16 or 4.1R<sup>60</sup>Δ16,18 isoforms were purified, loaded on to linear 5–20 %

sucrose gradients, centrifuged and fractionated as described in the Experimental section. Protein distribution along the gradient was analysed by immunoblotting using the 10b anti-4.1R antibody (Figure 1B) or an anti-GST antibody (results not shown). Proteins GST, GST-4.1R<sup>135</sup>Δ16 and GST-4.1R<sup>80</sup>Δ16 fractionated at the top of the gradient, whereas the small protein GST-4.1R<sup>60</sup>Δ16,18 did not remain at the top of the gradient but fractionated with a sedimentation coefficient higher than that of large 4.1R proteins. These results demonstrate the different sedimentation behaviour of the small protein GST-4.1R<sup>60</sup>Δ16,18, indicating that this protein is self-associated.

### Protein 4.1R<sup>60</sup>Δ16,18 binds both to itself and to other 4.1R<sup>60</sup> isoform types, but not to a 4.1R<sup>80</sup> isoform type

We next investigated whether 4.1R<sup>60</sup>Δ16,18 protein was able to establish intermolecular interactions with other 4.1R<sup>60</sup> proteins. Glutathione-Sepharose beads coupled with GST or GST-4.1R<sup>60</sup>Δ16,18 were incubated with <sup>35</sup>S-Met-labelled *in vitro*-translated 4.1R<sup>60</sup>Δ16 (Figure 2A). As revealed by autoradiographs, protein 4.1R<sup>60</sup>Δ16 bound to GST-4.1R<sup>60</sup>Δ16,18 (Figure 2A, lane 3), but not to GST control beads (Figure 2A, lane 2). This result confirmed that 4.1R<sup>60</sup> isoform types form intermolecular interactions. In addition, the fusion protein GST-4.1R<sup>60</sup>Δ16





**Figure 2** Protein 4.1R<sup>60</sup>Δ16,18 binds to itself and to another 4.1R<sup>60</sup> isoform, but not to a 4.1R<sup>80</sup> isoform

(A) Protein 4.1R<sup>60</sup>Δ16,18 binds to another 4.1R<sup>60</sup> isoform. <sup>35</sup>S-Met-labelled, *in vitro*-translated 4.1R<sup>60</sup>Δ16 protein (lane 1, input) was incubated with glutathione–Sephacryl beads coupled with GST (lane 2) or GST–4.1R<sup>60</sup>Δ16,18 (lane 3). The autoradiograph shows that the radiolabelled protein binds to GST–4.1R<sup>60</sup>Δ16,18. (B) Protein 4.1R<sup>80</sup>Δ16 does not interact with itself or with 4.1R<sup>60</sup>Δ16,18. Lysates from COS-7 cells expressing 4.1R<sup>80</sup>Δ16–GFP were incubated with glutathione–Sephacryl beads coupled with GST, GST–4.1R<sup>60</sup>Δ16,18 or GST–4.1R<sup>80</sup>Δ16. Immunoblots of the COS-7 lysate (lane 1, input), the bound fractions (lanes 2–4) and the unbound fractions (lanes 5–7) using the anti-GFP antibody are shown. Similarly, immunoblots of lysates from COS-7 cells expressing 4.1R<sup>60</sup>Δ16,18–GFP (lane 8, input) and incubated with GST (lane 9) or GST–4.1R<sup>60</sup>Δ16,18 (lane 10) are shown. (C) Size distribution of GST–4.1R<sup>60</sup>Δ16,18 analysed by size-exclusion chromatography, followed by immunoblotting with the anti-4.1R antibody, 10b. The positions of the protein standards thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa) and BSA (67 kDa) are indicated. (D) Sedimentation velocity data showing the oligomeric state of protein GST–4.1R<sup>60</sup>Δ16,18. Sedimentation velocity data were evaluated using a continuous *c(s)* distribution model. We show *c(s)* distribution at 11.5 °C.

behaved like GST–4.1R<sup>60</sup>Δ16,18 in sucrose-density-gradient centrifugation assays (results not shown).

To determine whether 4.1R<sup>60</sup> and 4.1R<sup>80</sup> isoform types establish intermolecular interactions, we performed pull-down assays using GST–4.1R<sup>60</sup>Δ16,18 bound to Sepharose beads on COS-7 lysates expressing isoform 4.1R<sup>80</sup>Δ16–GFP. As shown in Figure 2(B), these two proteins did not interact (lane 3). This Figure also shows that proteins 4.1R<sup>80</sup>Δ16–GFP and GST–4.1R<sup>80</sup>Δ16 are not associated (Figure 2B, lane 4), whereas the same type of experiments confirmed the interactions between proteins 4.1R<sup>60</sup>Δ16,18–GFP and GST–4.1R<sup>60</sup>Δ16,18 (Figure 2B, lane 10). Taken together, these results imply that 4.1R<sup>60</sup> isoforms are able to establish intermolecular interactions, while 4.1R<sup>80</sup> isoforms are not. They also indicate that 4.1R<sup>60</sup> and 4.1R<sup>80</sup> isoforms do not interact. These results are in close agreement with those from the sedimentation assays.

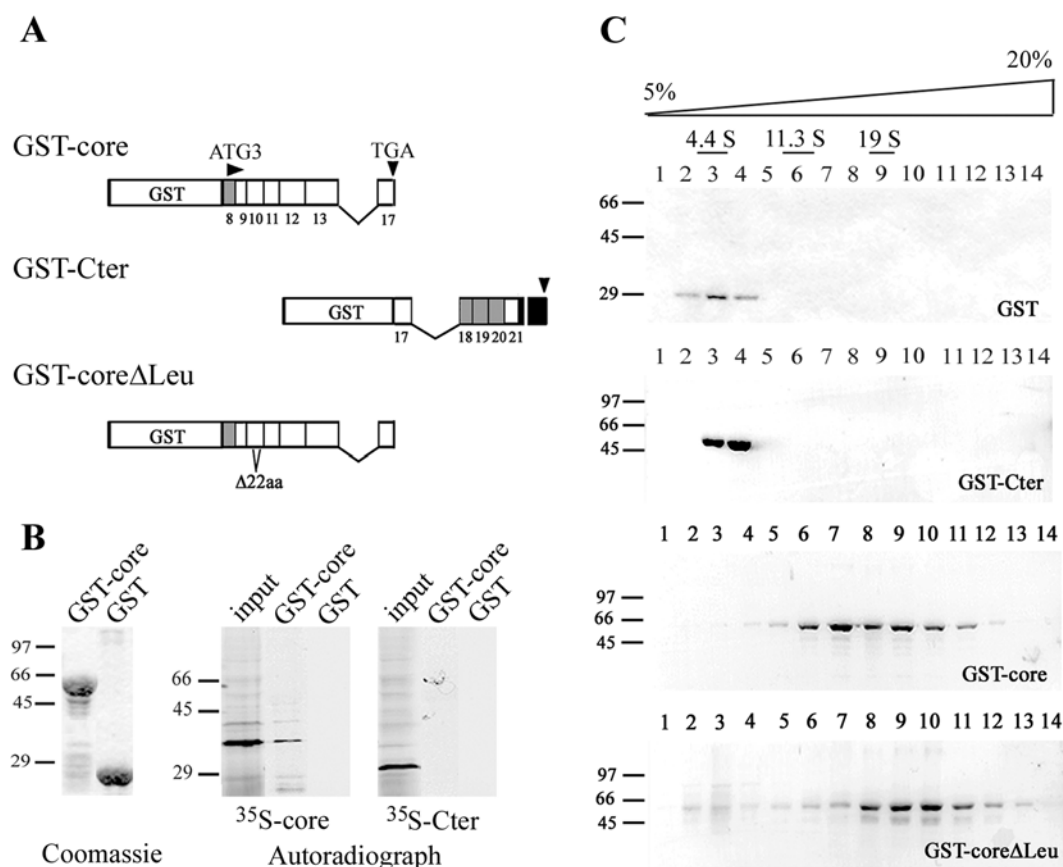
Protein 4.1R<sup>60</sup>Δ16,18 exhibits a relatively disperse distribution in the sucrose gradients, suggesting that it forms oligomers of more than two subunits. When analysed by size-exclusion chromatography, the purified GST–4.1R<sup>60</sup>Δ16,18 protein was eluted in several peak fractions covering a wide range of molecular sizes (Figure 2C). We observed a major peak at fractions 41–44 (which may correspond to the monomer and the dimer), a minor peak at fraction 37 (tetramer) and a third peak at fractions 29–32 (octamer). Further examination of GST–4.1R<sup>60</sup>Δ16,18 oligomer-

ization by sedimentation velocity analysis (Figure 2D) showed the presence of several species with sedimentation coefficients, corrected at 20 °C in water, of 6.97, 12.1, 19.7 and 26.4 S, consistent with the oligomeric forms detected by size-exclusion chromatography.

#### A conserved region is involved in 4.1R self-association

To identify the region involved in 4.1R<sup>60</sup> self-association, we divided the molecule into two halves, previously designated by us as the 'core' and 'Cter' regions, and each fragment was fused to GST (Figure 3A) [14,32]. Pull-down assays were performed using the recombinant protein GST–core bound to Sepharose beads and incubated with <sup>35</sup>S-Met-labelled, *in vitro*-translated core or Cter protein regions (Figure 3B). As revealed by the autoradiographs, <sup>35</sup>S-Met–core bound to GST–core but not to GST control beads (Figure 3B, autoradiograph, <sup>35</sup>S-core), while <sup>35</sup>S-Met–Cter bound neither to GST–core nor to GST (Figure 3B, autoradiograph, <sup>35</sup>S-Cter). These results indicate that the core region is involved in 4.1R self-association.

The involvement of the core region in 4.1R self-association was further analysed by sedimentation assays (Figure 3C). While GST and GST–Cter localized in the upper fractions of the gradient, GST–core distributed along the gradient with a peak in fractions 7–9, thus resembling the distribution of the full-length



**Figure 3** Identification of the 4.1R region involved in self-association

(A) Schematic representation of the exon map for the 4.1R cDNA constructs used in (B) and (C). The core and C-ter regions comprise amino acid residues 368–692 and 693–864 respectively (Swiss-Prot accession number P11171). (B) <sup>35</sup>S-Met-labelled, *in vitro*-translated proteins core and Cter (<sup>35</sup>S-core and <sup>35</sup>S-Cter) were incubated with glutathione–Sephadex beads coupled with GST or GST–core. Coomassie-stained gel showing the recombinant proteins (Coomassie) is shown. Autoradiograph of the <sup>35</sup>S-Met-labelled, *in vitro*-translated 4.1R proteins, inputs, and of the fractions bound to GST and GST–core (autoradiograph) is shown. <sup>35</sup>S-core but not <sup>35</sup>S-Cter binds to GST–core. (C) GST and the GST fusion proteins represented in (A) were purified and analysed by sedimentation assays as described in the Experimental section. Fractions were analysed by immunoblotting using the anti-GST antibody. The sucrose gradient, the fractions collected and the positions of sedimentation of the standard proteins, BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S), are all indicated. The core region, but not the Cter region, migrated with a high sedimentation coefficient, thus indicating its ability to self-associate. The leucine-zipper-like motif present in the FERM domain is not involved in self-association.

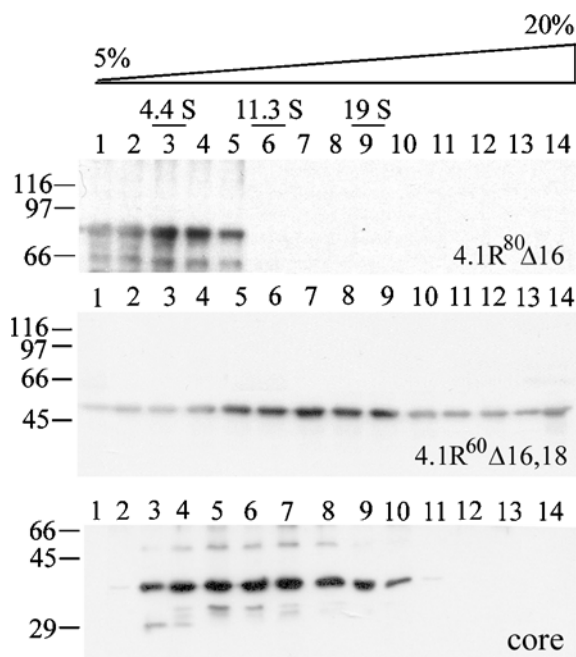
protein GST–4.1R<sup>60</sup>Δ16,18. These results confirm that the core region is responsible for 4.1R self-association.

As the core region contained 22 amino acids comprised of heptad repeats of leucine residues resembling a putative leucine zipper motif [14], we wondered whether they were involved in this intermolecular interaction. The behaviour of a fusion protein lacking the 22 amino acids and designated GST–coreΔLeu was analysed by sedimentation assays (Figure 3C, lower panel). The results show that while the core region was involved in 4.1R<sup>60</sup> self-association, the heptad repeats of leucine residues were not responsible for such interactions. Although the sedimentation assays revealed that GST alone localized in the upper fractions of the gradient (Figure 3C), we wanted to demonstrate that the fusion of GST to 4.1R isoforms does not artificially contribute to 4.1R self-association. The biochemical behaviour of 4.1R<sup>80</sup>Δ16, 4.1R<sup>60</sup>Δ16,18 and the core region, after removing GST (see the Experimental section), was analysed by sucrose-density-gradient centrifugation assays (Figure 4). Protein 4.1R<sup>80</sup>Δ16 fractionated at the top of the gradient, whereas the small 4.1R<sup>60</sup>Δ16,18 isoform and the core region fractionated with a high sedimentation coefficient. These results demonstrate that GST influences neither the self-association of the 4.1R<sup>60</sup> isoform and the core

region nor the differential sedimentation behaviour of 4.1R isoform types.

#### The 4.1R<sup>80</sup> isoform 4.1R<sup>80</sup>Δ16 was able to self-associate after denaturing–renaturing treatments

It was intriguing that all 4.1R isoforms contained the core region responsible for 4.1R<sup>60</sup> self-association, but that not all of them were able to self-associate. One possible explanation for this is that the core region was exposed in 4.1R<sup>60</sup>, but not in 4.1R<sup>80</sup> and 4.1R<sup>135</sup> isoforms. We reasoned that 4.1R<sup>80</sup> and 4.1R<sup>135</sup> isoforms would have the ability to self-associate if their core region were exposed. Thus we surmised that submitting the protein to denaturing conditions first and allowing it to renature slowly would favour the exposure of the core region, thus facilitating self-association before protein folding. To determine whether this indeed occurred, protein GST–4.1R<sup>80</sup>Δ16 was dialysed with a series of the appropriate buffer (see the Experimental section) containing consecutively greater concentrations of urea. The protein was then allowed to renature slowly by dialysing with a series of the same buffer containing consecutively lower concentrations of urea. The protein was loaded on to a 5–20 % sucrose-density gradient,



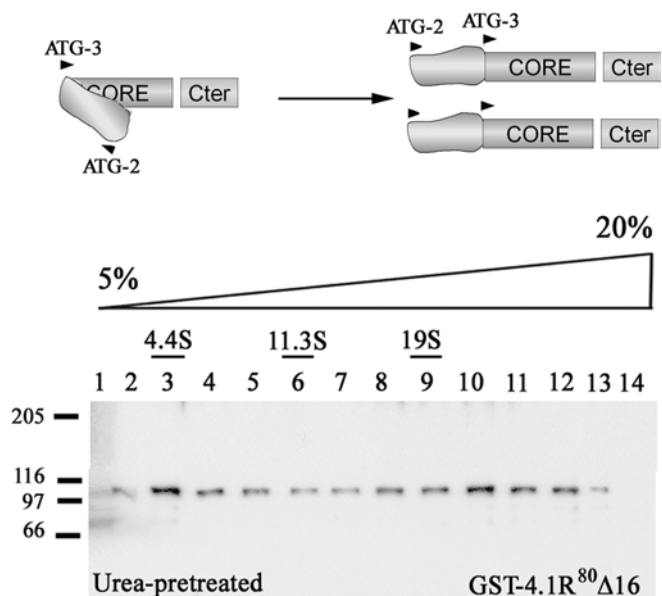
**Figure 4** Removal of the GST does not influence the self-association of protein 4.1R<sup>60</sup>

GST was removed from fusion proteins GST–4.1R<sup>80</sup>Δ16, GST–4.1R<sup>60</sup>Δ16,18 and GST–core as described in the Experimental section. Proteins 4.1R<sup>80</sup>Δ16, 4.1R<sup>60</sup>Δ16,18 and the core region were submitted to sedimentation assays and fractions collected were analysed by immunoblotting using the anti-4.1R 10b antibody to detect 4.1R<sup>80</sup>Δ16, 4.1R<sup>60</sup>Δ16,18 and the anti-4.1R 764 antibody to detect the core region. The sucrose gradient, the fractions collected and the positions of sedimentation of the standard proteins, BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S), are all indicated. 4.1R<sup>60</sup>Δ16,18 protein and the core region, but not 4.1R<sup>80</sup>Δ16, migrated with a high sedimentation coefficient, thus indicating their ability to self-associate independently of the GST.

centrifuged and fractionated as described in the Experimental section, and the results were analysed by immunoblotting using the 10b anti-4.1 antibody (Figure 5). It was observed that protein GST–4.1R<sup>80</sup>Δ16 distributed along the entire gradient with two peaks: the first is in fraction 3 and the other is in fraction 10, which may correspond to the monomeric and the self-associated species respectively. These results indicate that protein 4.1R<sup>80</sup>Δ16 is also able to self-associate, but that the core region of the molecule needs to be exposed for this interaction to occur. As it could be argued that the higher sedimenting species corresponded to aggregated protein that had not properly refolded, we performed sedimentation velocity analysis with protein GST–4.1R<sup>80</sup>Δ16 after denaturation–renaturation treatments. We detected species with sedimentation coefficients, corrected at 20 °C in water, of 6.5, 10 and 15.9 S (results not shown), which may correspond to the monomeric, dimeric and tetrameric forms. These results indicate that protein GST–4.1R<sup>80</sup>Δ16 has the ability to oligomerize.

#### The N-terminal region of the FERM domain interacts with the core region

The 4.1R FERM domain has been shown to have a cloverleaf-like crystal structure [43]. One major difference between the 4.1R<sup>80</sup> and 4.1R<sup>60</sup> isoform types is that the former contains a complete FERM domain, while the latter lacks the N-terminal end of this domain, the N-lobe and most of the  $\alpha$ -lobe (Figure 6A and 6B). It is reasonable to hypothesize that the N-terminal region of the FERM domain present in 4.1R<sup>80</sup> proteins masks the core



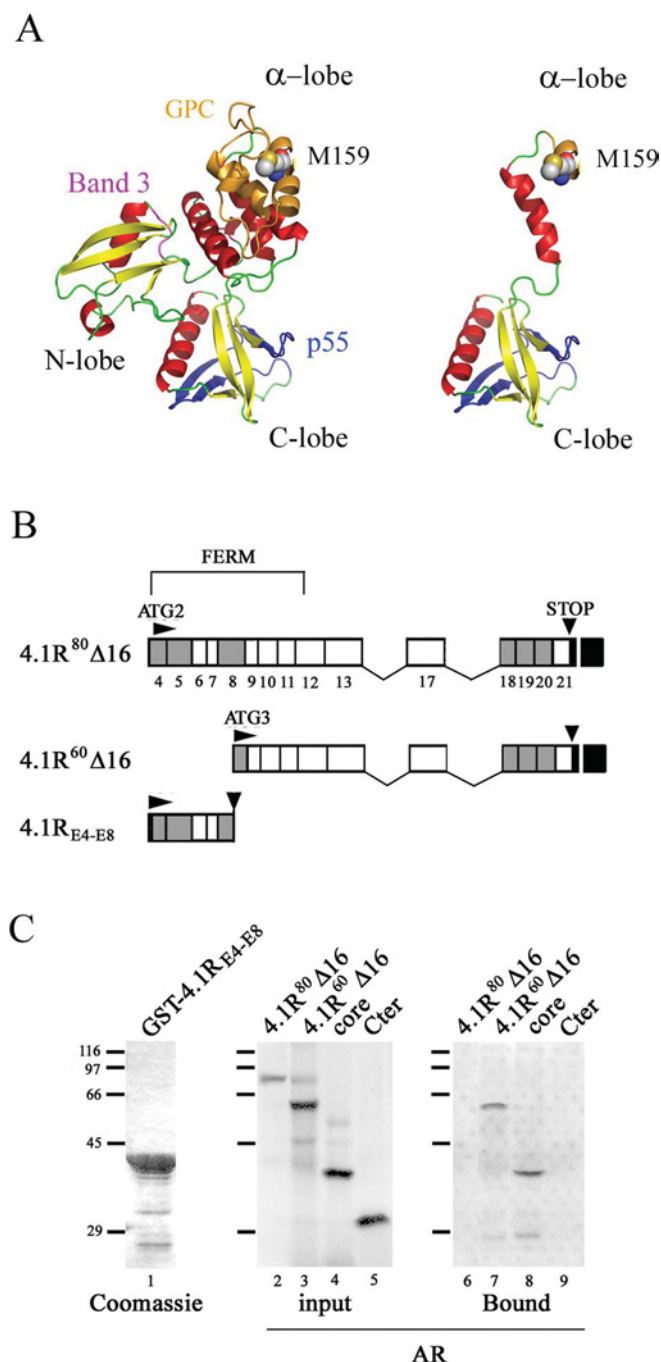
**Figure 5** Protein 4.1R<sup>80</sup>Δ16 self-association is achieved after denaturing–renaturing treatments

The fusion protein GST–4.1R<sup>80</sup>Δ16 was dialysed in the presence of urea as described in the Experimental section, and slowly renatured by removing the denaturing agent from the solution before being loaded on to the gradient. Immunoblot using the anti-4.1R antibody 10b revealed that GST–4.1R<sup>80</sup>Δ16 protein is then not only distributed in the upper part of the gradient. The sucrose gradient, the fractions collected and the positions of sedimentation of the standard proteins, BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S), are all indicated.

region in this type of isoforms. To investigate this, we performed pull-down assays using a GST fusion protein containing the 158 N-terminal amino acids of the FERM domain that are absent from 4.1R<sup>60</sup> isoforms (GST–4.1R<sub>E4-E8</sub>). The recombinant protein GST–4.1R<sub>E4-E8</sub> bound to Sepharose beads (Figure 6C, lane 1, Coomassie) was incubated with <sup>35</sup>S-Met-labelled *in vitro*-translated 4.1R<sup>80</sup>Δ16 protein, 4.1R<sup>60</sup>Δ16 protein, the core region or the Cter region (Figure 6C, inputs). As revealed by the autoradiographs, GST–4.1R<sub>E4-E8</sub> bound the small 4.1R<sup>60</sup>Δ16 isoform and the core region (Figure 6C, lanes 7 and 8), but not the large 4.1R<sup>80</sup>Δ16 isoform or the Cter region (Figure 6C, lanes 6 and 9). These results support the hypothesis that the core region is hidden in the large 4.1R isoforms by the intramolecular interaction between the N-terminus of the FERM domain and the core region.

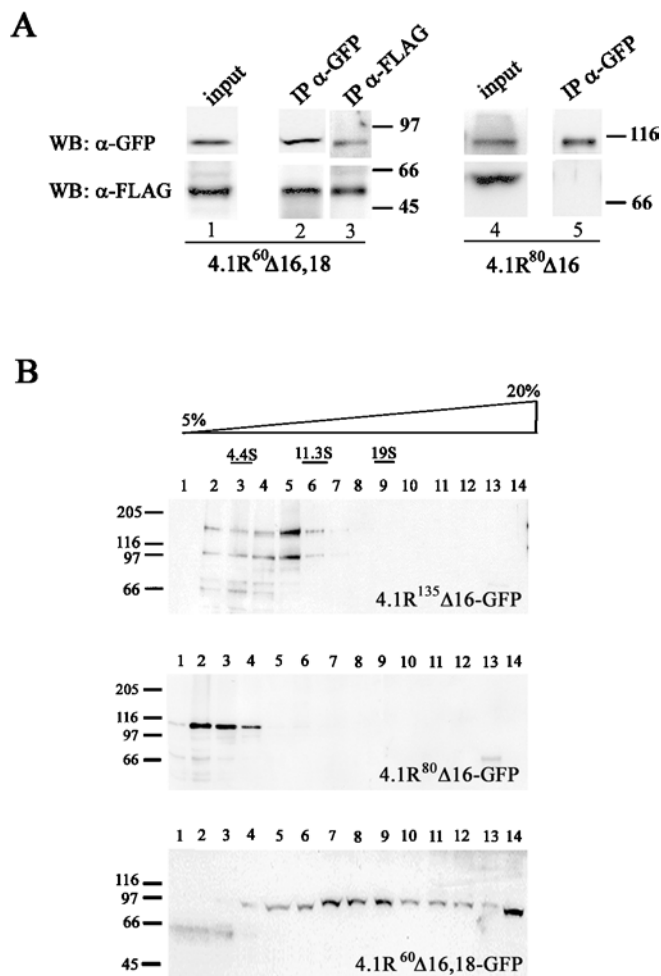
#### 4.1R<sup>60</sup> isoforms, but not those of 4.1R<sup>80</sup>, are isolated within the same complex *in vivo*

The results reported above indicate that small 4.1R<sup>60</sup> isoforms self-associate *in vitro*. We next investigated whether multiple 4.1R<sup>60</sup>Δ16,18 molecules could be isolated within the same complex *in vivo*. For this purpose we co-expressed proteins 4.1R<sup>60</sup>Δ16,18 tagged with two different tags, one at the N-terminus, FLAG–4.1R<sup>60</sup>Δ16,18, and one at the C-terminus, 4.1R<sup>60</sup>Δ16,18–GFP, and performed co-immunoprecipitation experiments with the appropriate antibodies (Figure 7A). Immunoprecipitation with the anti-GFP antibody, followed by immunoblotting with the anti-FLAG antibody, revealed the co-immunoprecipitation of the two proteins from lysates of co-transfected COS-7 cells (Figure 7A, lane 2). When the anti-FLAG antibody was used for the immunoprecipitation assay and the anti-GFP antibody for the immunoblotting, both proteins also co-immunoprecipitated (Figure 7A, lane 3). In contrast with these results, when the 4.1R<sup>80</sup>



**Figure 6** The N-terminal sequence of the FERM domain interacts with the core region

(A) Left: ribbon diagram of the FERM domain of 4.1R [43]. The crystal structure was taken from PDB 1GG3.  $\alpha$ -Helices are coloured red,  $\beta$ -sheets yellow and loops green. Binding sites for band 3, glycophorin C (GPC) and p55 are represented in purple, orange and blue respectively. The position of the initial methionine of 4.1R<sup>60</sup> isoforms in the  $\alpha$ -lobe, residue Met<sup>159</sup>, is displayed in space-filling representation. The right-hand side shows the FERM domain that remains in 4.1R<sup>60</sup> isoforms. Modelling was performed using the PyMol program (<http://www.pymol.org>). (B) Schematic representations of the exon maps for 4.1R<sup>80</sup> $\Delta$ 16 and 4.1R<sup>60</sup> $\Delta$ 16 showing the additional N-terminal extension (4.1R<sub>E4-E8</sub>) of the FERM domain present in protein 4.1R<sup>80</sup> $\Delta$ 16. (C) Pull-down *in vitro* assays. <sup>35</sup>S-Met-labelled, *in vitro*-translated proteins 4.1R<sup>80</sup> $\Delta$ 16 and 4.1R<sup>60</sup> $\Delta$ 16, the core region and the Cter region (lanes 2–5, input) were incubated with glutathione–Sepharose beads coupled with GST–4.1R<sub>E4-E8</sub> (lane 1, Coomassie). The autoradiographs of the bound fractions (lanes 6–9, Bound) show that the N-terminal of the FERM domain (GST–4.1R<sub>E4-E8</sub>) binds to 4.1R<sup>80</sup> $\Delta$ 16 protein and the core region, but not to 4.1R<sup>60</sup> $\Delta$ 16 or to the Cter region.



**Figure 7** The small 4.1R<sup>60</sup> $\Delta$ 16,18 isoform is self-associated *in vivo*

(A) 4.1R<sup>60</sup> isoforms, but not those of 4.1R<sup>80</sup>, are isolated within the same complex. COS-7 cells were co-transfected to co-express proteins 4.1R<sup>60</sup> $\Delta$ 16,18–GFP and FLAG–4.1R<sup>60</sup> $\Delta$ 16,18 or proteins 4.1R<sup>80</sup> $\Delta$ 16–GFP and FLAG–4.1R<sup>80</sup> $\Delta$ 16. Cell lysates (lanes 1 and 4, input) were immunoprecipitated (IP) with the polyclonal antibodies anti-GFP (lanes 2 and 5) or anti-FLAG (lane 3). Samples were analysed by immunoblotting using the monoclonal antibodies anti-GFP (WB:  $\alpha$ -GFP), and anti-FLAG (WB:  $\alpha$ -FLAG). (B) Lysates of cells transfected with cDNAs encoding 4.1R<sup>135</sup> $\Delta$ 16–GFP, 4.1R<sup>80</sup> $\Delta$ 16–GFP or 4.1R<sup>60</sup> $\Delta$ 16,18–GFP were loaded on to 5–20% sucrose-density gradients, centrifuged and fractionated as described in the Experimental section. Fractions were analysed by immunoblotting using an anti-GFP antibody. The sucrose-density gradient, the fractions collected and the position of sedimentation of the standard proteins, BSA (4.4 S), catalase (11.3 S) and thyroglobulin (19 S), are all indicated. Positions of molecular-mass markers (kDa) are shown on the left. Note that the smallest protein, 4.1R<sup>60</sup> $\Delta$ 16,18–GFP (lower panel), migrates with a higher sedimentation coefficient than that of proteins 4.1R<sup>135</sup> $\Delta$ 16–GFP and 4.1R<sup>80</sup> $\Delta$ 16–GFP.

isoform 4.1R<sup>80</sup> $\Delta$ 16 was co-expressed with the two different tags, FLAG–4.1R<sup>80</sup> $\Delta$ 16 and 4.1R<sup>80</sup> $\Delta$ 16–GFP, and assayed in co-immunoprecipitation experiments, they did not co-immunoprecipitate (Figure 7A, lane 5). These results are in agreement with those obtained from the pull-down assays (Figure 2).

#### Isoforms of protein 4.1R exogenously expressed in COS-7 cells show different sedimentation behaviours

We next performed sucrose-density-gradient centrifugation assays using extracts from COS-7 cells transfected with 4.1R<sup>135</sup> $\Delta$ 16–GFP, 4.1R<sup>80</sup> $\Delta$ 16–GFP or 4.1R<sup>60</sup> $\Delta$ 16,18–GFP cDNAs to compare the sedimentation behaviour of the expressed proteins

with that of the recombinant proteins analysed in Figure 1. The lysates were loaded on to linear 5–20 % sucrose-density gradients, centrifuged and fractionated as described in the Experimental section. The distribution of the proteins along the gradient was determined by immunoblotting using an anti-GFP antibody (Figure 7B). We observed that proteins 4.1R<sup>135</sup>Δ16–GFP and 4.1R<sup>80</sup>Δ16–GFP remained at the top of the gradient, whereas the small protein 4.1R<sup>60</sup>Δ16,18–GFP did not, but instead fractionated with a sedimentation coefficient higher than that of the large 4.1R proteins. Similar distribution patterns were observed for the 4.1R<sup>135</sup> isoforms 4.1R<sup>135</sup>Δ16,19 and 4.1R<sup>135</sup>Δ16,18,19 and the 4.1R<sup>80</sup> isoforms 4.1R<sup>80</sup>Δ16,18 and 4.1R<sup>80</sup>Δ16,19 (results not shown). The similar sedimentation behaviour observed for isoform 4.1R<sup>60</sup>Δ16,18 isolated from COS-7 cells and for the recombinant protein GST–4.1R<sup>60</sup>Δ16,18 supports the notion that the small 4.1R<sup>60</sup>Δ16,18 isoform also self-associates *in vivo*.

## DISCUSSION

Protein self-association is a very common phenomenon that can confer several structural and functional advantages on proteins, including improved stability, control over the accessibility and specificity of active sites, and increased complexity. Recent structural and biophysical studies show that protein self-association is a key factor in the regulation of proteins such as enzymes, ion channels, receptors and transcription factors [44]. The present study shows that protein 4.1R can also be included within the growing list of proteins that are known to be capable of self-association. The sedimentation data indicate that isoform 4.1R<sup>60</sup>Δ16,18, previously isolated from human T-cells and whose sequence is contained in 4.1R<sup>135</sup> and 4.1R<sup>80</sup> isoforms [32], oligomerizes. Pull-down assays confirmed that 4.1R<sup>60</sup>Δ16,18 interacts with itself, but not with a larger 4.1R<sup>80</sup> isoform. Interestingly, a region conserved in all 4.1R isoforms, previously designated by us as the core region [32], is involved in self-association.

The fact that all isoforms of protein 4.1R contained the core region and only the small isoform 4.1R<sup>60</sup>Δ16,18 was found to be self-associated suggested to us that the core region must be hidden in 4.1R<sup>80</sup> and 4.1R<sup>135</sup> isoform types, but not in 4.1R<sup>60</sup> isoforms. This idea was supported by the results of denaturing–renaturing experiments in which an initially non-self-associated 4.1R<sup>80</sup> isoform became self-associated (Figure 5). Thus large isoforms of protein 4.1R have the ability to self-associate, but the core region of the molecule needs to become exposed for this interaction to occur.

The core region comprises the N-terminal end of 4.1R<sup>60</sup> isoforms; however, up to 367 and 158 amino acids are added N-terminally to this region in the 4.1R<sup>135</sup> and 4.1R<sup>80</sup> isoforms respectively. It is reasonable to propose that these extra N-terminal extensions might be responsible for hiding the core region. This hypothesis was supported by the results from the pull-down assays showing that the N-terminal region of the FERM domain, which was absent from 4.1R<sup>60</sup> isoforms but present in the extensions of 4.1R<sup>135</sup> and 4.1R<sup>80</sup> isoforms, specifically bound to the core region (Figure 6). It is of note that hiding of the core region was specific to the 4.1R amino acid sequence, as adding an extra N-terminal extension comprised of the GST sequence to that of the small protein 4.1R<sup>60</sup>Δ16,18 did not prevent self-association of the fusion protein (Figure 1). Taken together, our results imply that 4.1R<sup>60</sup> isoforms, lacking the N-terminal region of the FERM domain, have exposed the core region and are therefore self-associated. By contrast, self-association of proteins 4.1R containing a complete FERM domain appears to be modulated by intramolecular interactions.

It is noteworthy that 4.1R<sup>60</sup> isoforms are the smallest 4.1R isoforms and that the presence or absence of the N-terminal extension of 4.1R<sup>135</sup> and 4.1R<sup>80</sup> must confer special characteristics to the different 4.1R isoforms. The results presented here are consistent with those of our previous studies of the nuclear/cytoplasmic distribution of 4.1R proteins, wherein we observed that 4.1R<sup>60</sup> isoforms were predominantly distributed to the nucleus due to the nuclear targeting effect of the core region [32]. By contrast, 4.1R<sup>135</sup> isoforms also contain the core region and were predominantly localized at non-nuclear sites, implying that the core region is hidden by their extra N-terminal domain [36,41].

A mechanism preventing protein self-association through intramolecular interactions has been reported for many proteins, such as the ERM (ezrin, radixin, moesin) family of proteins belonging to the band 4.1 superfamily. Indeed, monomeric ezrin possesses a conformationally hidden C-terminal domain that, when exposed, can bind to an N-terminal domain (the FERM domain) of a second molecule [45]. It is noteworthy that while two different domains, the N- and C-termini, are involved in ERM intermolecular interactions, only one region, the core region, is involved in 4.1R intermolecular interactions. The WASP (Wiskott–Aldrich syndrome protein) family and the formins mDia1 (a mammalian homologue of the *Drosophila* diaphanous protein) and FHOD1 (forming homology 2 domain containing protein) also form intramolecular interactions in order to hide regions involved in protein interactions [46,47]. These designs virtually ensure that self-association is physiologically important and imply that self-interacting protein complexes have capabilities that monomers do not. Self-association of spectrin, one of the major structural erythrocyte proteins, proved to be essential for normal erythroid shape and mechanical stability [48]. Mutations in spectrin that impair its ability to self-associate lead to clinically significant forms of HE (hereditary elliptocytosis) and HPP (hereditary pyropoikilocytosis) [49]. Future investigation will shed light on the functional consequences of protein 4.1R self-association.

We thank Dr M. A. Alonso (Centro de Biología Molecular 'Severo Ochoa', Madrid, Spain) for invaluable discussions. We are very grateful to Dr C. Alfonso (Centro de Investigaciones Biológicas, Madrid, Spain) for advice and assistance on sedimentation velocity assays. We also thank A. Gosálbez for her excellent technical support. This work was supported by grant number BFU2005-01825 from the Ministerio de Educación y Ciencia, Spain. C.M.P.-F. and E.L. were post-doctoral and pre-doctoral fellows of the Ministerio de Educación y Ciencia, Spain.

## REFERENCES

- 1 An, X. L., Takakuwa, Y., Manno, S., Han, B. J., Gascard, P. and Mohandas, N. (2001) Structural and functional characterization of protein 4.1R-phosphatidylserine interaction: potential role in 4.1R sorting within cells. *J. Biol. Chem.* **276**, 35778–35785
- 2 An, X., Zhang, X., Debnath, G., Baines, A. J. and Mohandas, N. (2006) Phosphatidylinositol-4,5-bisphosphate (PIP2) differentially regulates the interaction of human erythrocyte protein 4.1 (4.1R) with membrane proteins. *Biochemistry* **45**, 5725–5732
- 3 Conboy, J. G. (1993) Structure, function, and molecular genetics of erythroid membrane skeletal protein 4.1 in normal and abnormal red blood cells. *Semin. Hematol.* **30**, 58–73
- 4 Anderson, R. A., Correas, I., Mazzucco, C., Castle, J. D. and Marchesi, V. T. (1988) Tissue-specific analogues of erythrocyte protein 4.1 retain functional domains. *J. Cell Biochem.* **37**, 269–284
- 5 Granger, B. L. and Lazarides, E. (1984) Membrane skeletal protein 4.1 of avian erythrocytes is composed of multiple variants that exhibit tissue-specific expression. *Cell* **37**, 595–607
- 6 Cohen, C. M., Foley, S. F. and Korsgren, C. (1982) A protein immunologically related to erythrocyte band 4.1 is found on stress fibres on non-erythroid cells. *Nature* **299**, 648–650

- 7 Leto, T. L., Pratt, B. M. and Madri, J. A. (1986) Mechanisms of cytoskeletal regulation: modulation of aortic endothelial cell protein band 4.1 by the extracellular matrix. *J. Cell Physiol.* **127**, 423–431
- 8 Correas, I. (1991) Characterization of isoforms of protein 4.1 present in the nucleus. *Biochem. J.* **279**, 581–585
- 9 Krauss, S. W., Chasis, J. A., Rogers, C., Mohandas, N., Krockmalnic, G. and Penman, S. (1997) Structural protein 4.1 is located in mammalian centrosomes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7297–7302
- 10 Cohen, A. R., Woods, D. F., Marfatia, S. M., Walther, Z., Chishti, A. H., Anderson, J. M. and Wood, D. F. (1998) Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *J. Cell Biol.* **142**, 129–138
- 11 Mattagajasingh, S. N., Huang, S. C., Hartenstein, J. S. and Benz, Jr, E. J. (2000) Characterization of the interaction between protein 4.1R and ZO-2. A possible link between the tight junction and the actin cytoskeleton. *J. Biol. Chem.* **275**, 30573–30585
- 12 Kontogianni-Konstantopoulos, A., Huang, S. C. and Benz, Jr, E. J. (2000) A nonerythroid isoform of protein 4.1R interacts with components of the contractile apparatus in skeletal myofibers. *Mol. Biol. Cell* **11**, 3805–3817
- 13 Mattagajasingh, S. N., Huang, S. C., Hartenstein, J. S., Snyder, M., Marchesi, V. T. and Benz, E. J. (1999) A nonerythroid isoform of protein 4.1R interacts with the Nuclear Mitotic Apparatus (NuMA) protein. *J. Cell Biol.* **145**, 29–43
- 14 Perez-Ferreiro, C. M., Luque, C. M. and Correas, I. (2001) 4.1R proteins associate with interphase microtubules in human T cells: a 4.1R constitutive region is involved in tubulin binding. *J. Biol. Chem.* **276**, 44785–44791
- 15 Perez-Ferreiro, C. M., Vernos, I. and Correas, I. (2004) Protein 4.1R regulate interphase microtubule organization at the centrosome. *J. Cell Sci.* **117**, 6197–6206
- 16 Huang, S. C., Jagadeeswaran, R., Liu, E. S. and Benz, Jr, E. J. (2004) Protein 4.1R, a microtubule-associated protein involved in microtubule aster assembly in mammalian mitotic extract. *J. Biol. Chem.* **279**, 34595–34602
- 17 Krauss, S. W., Lee, G., Chasis, J. A., Mohandas, N. and Heald, R. (2004) Two protein 4.1 domains essential for mitotic spindle and aster microtubule dynamics and organization *in vitro*. *J. Biol. Chem.* **279**, 27591–27598
- 18 De Cárcer, G., Lallena, M. J. and Correas, I. (1995) Protein 4.1 is a component of the nuclear matrix of mammalian cells. *Biochem. J.* **312**, 871–877
- 19 Krauss, S. W., Heald, R., Lee, G., Nunomura, W., Gimm, J. A., Mohandas, N. and Chasis, J. A. (2002) Two distinct domains of protein 4.1 critical for assembly of functional nuclei *in vitro*. *J. Biol. Chem.* **277**, 44339–44346
- 20 Lallena, M. J., Martínez, C., Válcárcel, J. and Correas, I. (1998) Functional association of nuclear protein 4.1 with pre-mRNA splicing factors. *J. Cell Sci.* **111**, 1963–1971
- 21 Robb, V. A., Li, W., Gascard, P., Perry, A., Mohandas, N. and Gutmann, D. H. (2003) Identification of a third protein 4.1 tumor suppressor, Protein 4.1R, in meningioma pathogenesis. *Neurobiol. Dis.* **13**, 191–202
- 22 Conboy, J. (1999) The role of alternative pre-mRNA splicing in regulating the structure and function of skeletal protein 4.1. *Proc. Soc. Exp. Biol. Med.* **220**, 73–78
- 23 Tang, T. K., Qin, Z., Leto, T., Marchesi, V. T. and Benz, Jr, E. J. (1990) Heterogeneity of mRNA and protein products arising from the protein 4.1 gene in erythroid and nonerythroid tissues. *J. Cell Biol.* **110**, 617–624
- 24 Tang, T. K., Leto, T. L., Correas, I., Alonso, M. A., Marchesi, V. T. and Benz, Jr, E. J. (1988) Selective expression of an erythroid-specific isoform of protein 4.1. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3713–3717
- 25 Yang, G., Huang, S. C., Wu, J. Y. and Benz, Jr, E. J. (2005) An erythroid differentiation-specific splicing switch in protein 4.1R mediated by the interaction of SF2/ASF with an exonic splicing enhancer. *Blood* **105**, 2146–2153
- 26 Schischmanoff, P. O., Yaswen, P., Parra, M. K., Lee, G., Chasis, J. A., Mohandas, N. and Conboy, J. G. (1997) Cell shape-dependent regulation of protein 4.1 alternative pre-mRNA splicing in mammary epithelial cells. *J. Biol. Chem.* **272**, 10254–10259
- 27 Hou, V. C. and Conboy, J. G. (2001) Regulation of alternative pre-mRNA splicing during erythroid differentiation. *Curr. Opin. Hematol.* **8**, 74–79
- 28 Conboy, J. G., Chan, J. Y., Chasis, J. A., Kan, Y. W. and Mohandas, N. (1991) Tissue- and development-specific alternative RNA splicing regulates expression of multiple isoforms of erythroid membrane protein 4.1. *J. Biol. Chem.* **266**, 8273–8280
- 29 Chasis, J. A., Coulombel, L., Conboy, J., McGee, S., Andrews, K., Kan, Y. W. and Mohandas, N. (1993) Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. *J. Clin. Invest.* **91**, 329–338
- 30 Baklouti, F., Huang, S. C., Vulliamy, T. J., Delaunay, J. and Benz, Jr, E. J. (1997) Organization of the human protein 4.1 genomic locus: new insights into the tissue-specific alternative splicing of the pre-mRNA. *Genomics* **39**, 289–302
- 31 Gascard, P., Lee, G., Coulombel, L., Auffray, I., Lum, M., Parra, M., Conboy, J. G., Mohandas, N. and Chasis, J. A. (1998) Characterization of multiple isoforms of protein 4.1R expressed during erythroid terminal differentiation. *Blood* **92**, 4404–4414
- 32 Luque, C. M. and Correas, I. (2000) A constitutive region is responsible for nuclear targeting of 4.1R: modulation by alternative sequences results in differential intracellular localization. *J. Cell Sci.* **113**, 2485–2495
- 33 Leto, T. L. and Marchesi, V. T. (1984) A structural model of human erythrocyte protein 4.1. *J. Biol. Chem.* **259**, 4603–4608
- 34 Sun, C. X., Robb, V. A. and Gutmann, D. H. (2002) Protein 4.1 tumor suppressors: getting a FERMAP13 on growth regulation. *J. Cell Sci.* **115**, 3991–4000
- 35 Chishti, A. H., Kim, A. C., Marfatia, S. M., Lutchman, M., Hanspal, M., Jindal, H., Liu, S. C., Low, P. S., Rouleau, G., Mohandas, N. et al. (1998) The FERMAP domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane. *Trends Biochem. Sci.* **23**, 281–282
- 36 Luque, C. M., Lallena, M. J., Pérez-Ferreiro, C. M., de Isidro, Y., De Cárcer, G., Alonso, M. A. and Correas, I. (1999) The N-terminal 209-aa domain of high molecular-weight 4.1R isoforms abrogates 4.1R targeting to the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14925–14930
- 37 Luque, C. M., Perez-Ferreiro, C. M., Perez-Gonzalez, A., Englmeier, L., Koffa, M. D. and Correas, I. (2003) An alternative domain containing a leucine-rich sequence regulates nuclear cytoplasmic localization of protein 4.1R. *J. Biol. Chem.* **278**, 2686–2691
- 38 Correas, I., Speicher, D. W. and Marchesi, V. T. (1986) Structure of the spectrin-actin binding site of erythrocyte protein 4.1. *J. Biol. Chem.* **261**, 13362–13366
- 39 Lallena, M. J. and Correas, I. (1997) Transcription-dependent redistribution of nuclear protein 4.1 to SC35-enriched nuclear domains. *J. Cell Sci.* **110**, 239–247
- 40 Echeverri, C. J., Paschal, B. M., Vaughan, K. T. and Vallee, R. B. (1996) Molecular characterization of the 50-kD subunit of dynactin reveals function for the complex in chromosome alignment and spindle organization during mitosis. *J. Cell Biol.* **132**, 617–633
- 41 Luque, C. M., Lallena, M. J., Alonso, M. A. and Correas, I. (1998) An alternative domain determines nuclear localization in multifunctional protein 4.1. *J. Biol. Chem.* **273**, 11643–11649
- 42 Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J. and Schubert, D. (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. *Biophys. J.* **82**, 1096–1111
- 43 Han, B. G., Nunomura, W., Takakuwa, Y., Mohandas, N. and Jap, B. K. (2000) Protein 4.1R core domain structure and insights into regulation of cytoskeletal organization. *Nat. Struct. Biol.* **7**, 871–875
- 44 Marianayagam, N. J., Sunde, M. and Matthews, J. M. (2004) The power of two: protein dimerization in biology. *Trends Biochem. Sci.* **29**, 618–625
- 45 Louvet-Vallee, S. (2000) ERM proteins: from cellular architecture to cell signaling. *Biol. Cell* **92**, 305–316
- 46 Millard, T. H., Sharp, S. J. and Machesky, L. M. (2004) Signalling to actin assembly via the WASP (Wiskott-Aldrich syndrome protein)-family proteins and the Arp2/3 complex. *Biochem. J.* **380**, 1–17
- 47 Wallar, B. J. and Alberts, A. S. (2003) The formins: active scaffolds that remodel the cytoskeleton. *Trends Cell Biol.* **13**, 435–446
- 48 An, X., Lecomte, M. C., Chasis, J. A., Mohandas, N. and Gratzer, W. (2002) Share-response of the spectrin dimer-tetramer equilibrium in the red blood cell membrane. *J. Biol. Chem.* **277**, 31796–31800
- 49 Zhang, Z., Weed, S. A., Gallagher, P. G. and Morrow, J. S. (2001) Dynamic molecular modeling of pathogenic mutations in the spectrin self-association domain. *Blood* **98**, 1645–1653

Received 2 May 2006/19 July 2006; accepted 2 August 2006

Published as BJ Immediate Publication 2 August 2006, doi:10.1042/BJ20060644

